



Tilting at Windmills

The City University of New York's ACT Writing Exam

BRUCE CHADWICK

Kingsborough Community College

This article examines the CUNY-ACT as a high-stakes, standardized exit exam for developmental writing students at one CUNY school, Kingsborough Community College. I chart the political conditions at CUNY that led to the establishment of the exam, and its disruption of the existing assessment procedures already in place at Kingsborough. I present examples of ACT prompts and explain the test preparation course for students who have failed the exam numerous times. I critique the report that presents the rationale and procedure of testing presented to Kingsborough by the central office of CUNY, consisting of the CUNY Board of Trustees and Chancellor, in conjunction with New York City politicians, including the then mayor, Rudolph W. Giuliani. I explain the role of the CUNY central office in forcing the ACT to be implemented at Kingsborough without consideration of well-established research on validity in the area of writing assessment. I explain the destructive effects that the ACT has on Kingsborough students, especially on those who are non-native speakers and writers. I argue for better assessment procedures at Kingsborough derived from research in the area of writing assessment, and ask for greater coordination of effort among Kingsborough students, faculty, and the CUNY Central administration to establish an assessment policy that rests upon appropriate pedagogical practice and sound validity.

Bruce Chadwick is an adjunct assistant professor at Kingsborough Community College, CUNY, where he teaches courses in composition and developmental writing. He also teaches courses in composition at Long Island University, Brooklyn, where he chairs the Assessment Subcommittee. He has previously published articles on reading, grammar, and has contributed poetry and prose for Long Island University's faculty-student journal, *Downtown Brooklyn*.

Direct all correspondence to: Bruce Chadwick, Dept. of English, Kingsborough Community College CUNY, 2001 Oriental Blvd., Brooklyn, NY 11235-2398.

The Day of the Test

On Thursday, June 2, 2005, at 1:30 p.m., my basic writing students at Kingsborough Community College, a 2-year college in the City of New York University (CUNY) system, took the CUNY-ACT writing exam for which I had helped them prepare since the semester began on March 1. They joined some 300 other students in a large amphitheater where officials from the Testing Office administered the 60-minute timed test. This is standard operating procedure for every semester the test is administered. No teachers are allowed into the amphitheater during the test. The students are strictly timed. They are nervous, anxious; some will freeze up and not be able to write. Others will write but not be able to finish. If some students persist in trying to finish beyond the 60-minute time limit, their tests are confiscated. The students are not allowed to use dictionaries.

One week later I received the results of the test. Of the 22 students in my class who took the test, 10 passed. The 12 students who did not pass would have to repeat this course until they eventually passed. How did my students arrive at this point in their education, where they found themselves on a treadmill of test-taking? How and why did the college allow them to endure this experience? Why did the college turn to a standardized test as a way to assess writing? The experience of teaching this test preparation course and the unsatisfactory results from students taking it have prompted me to address these questions in this article.

The push for more standardized testing and assessment is not new. Since the late 1970s, state governments seem to have developed the notion that a uniform test, whether in multiple-choice form or in a simplified essay form, will ensure that students are learning. States such as California, Texas, and Florida especially, have viewed public education as an industry, where the emphasis is on results. In discussing the situation in Florida, Brossell (1996) noted, "Florida's government touts public education. If students are passing state tests in large numbers, they must be learning something in school. And if schools are demonstrably successful in teaching students, citizens are more likely to feel that they are getting a decent bang for their tax bucks and less likely to hassle their state legislators. This is . . . political reality" (p. 29). In Florida, Brossell reported that college students must take the College-Level Academic Skills Test (CLAST), "a check on basic reading, writing, and math abilities . . ." (p. 25). These testing programs, according to Florida's state leaders, "would help ensure that decent standards of academic competence would be upheld across the spectrum of Florida's schools and colleges" (p. 25). Brossell called the effect of CLAST on students "pernicious" (p. 28). Other states as well, have mandated similar standardized tests with similar consequences (see Fraizer, 2003). The situation at Kingsborough was not unique in this regard. In what follows, I describe the CUNY-ACT test's pernicious effect on the students and faculty at Kingsborough.

Background and Demographics

The majority of students I teach in this test-preparation course, called English W, are English-language learners (ELLs). They have arrived here either directly from their country of origin, or are second-generation children of

parents who arrived in Brooklyn to begin a new life. The students have entered Kingsborough Community College to begin their educational careers in such fields as business administration, travel and tourism, computer science, nursing, and in some cases, "liberal arts." Some are intent on transferring to 4-year colleges. Their command of spoken and written English varies, but for the most part they have great difficulty writing in English because of unfamiliarity with English syntax, grammar, and vocabulary. The students I teach have all passed their English as a Second Language (ESL) classes and have entered regular non-ESL remedial classes that they completed and passed. But they cannot move on to a regular, nonremedial freshman English course until they pass the CUNY-ACT. All remedial students, whether ESL or not, must take and pass Kingsborough's remedial courses before they can take English W, which is specifically for students who have passed everything but the CUNY-ACT test.

The students in a later English W course that I taught immediately after the one just noted reflect the difficulty of designing and teaching such a course. Twenty-five students registered for this course. Eighteen are students whose first language is not English; only seven students are first-language English speakers and writers. Of the 18 students, the breakdown is as follows: Seven are originally from China; one is from Korea; one from Bangladesh; one from Bosnia-Herzegovina; one from Albania; one from Israel (whose first language is Hebrew); one from the Ukraine; two from Russia; one from Puerto Rico; one from the Dominican Republic; and one from Mexico. This cultural and ethnic mix of students is typical for a Kingsborough class. All but three of these students failed the CUNY-ACT test more than once: Seven students failed the test twice; six failed the test three times; one student failed the test four times, and one student failed the test five times.

Although I know that not all my students will pass the CUNY-ACT, I am ever hopeful that at least half will pass. In the past, the average percentage of students in my test-prep courses who fail has been around 40% to 50%. Those who do not pass will feel their confidence and determination begin to slip away despite the fact that they work very hard. The students I teach are similar to the ones noted in Sternglass' (1997) book, *A Time to Know Them: A Longitudinal Study of Writing and Learning at the College Level*, in which she presented case studies of nine students at the City College of New York (CCNY), a 4-year branch of CUNY. She followed their lives and careers at CCNY and beyond, during the 1990s, when she had access to their work, and read all of their writing, contacting them frequently during their time at CCNY and after they had left. The students she wrote about shared with her their painful, often debilitating, struggles to succeed against great odds in their college careers. Most are the first in their families to go to college. She noted their growth in gradually adapting to disciplinary discourse, learning to write analytically, with power and fluency. These changes occurred over many years, after many trial-and-error attempts. The following is one of her general observations in conducting her study:

It would be possible to understand if sheer fatigue, economic pressures, and frustration at being unable to do the best work possible were to undermine the commitment of the students in this study to complete their academic responsibilities honorably

and with pride. But they did not falter—they made adjustments where necessary. They tried their best not to compromise their academic standards. Instead of just giving them credit for their success, the society should mobilize itself to improve the conditions under which such dedicated individuals can aspire to higher levels of personal and professional lives. Certainly better preparation at the precollege years is highly desirable, but in its absence in urban areas like New York City, the colleges should be set up so as to foster opportunities for students who earnestly wish to make up for their poor previous educational environments. (p. 106)

Sternglass realized the ambitions and determination of remedial students with whom she has worked. I, too, have seen how hard my students work to succeed, day in and day out, despite their difficulties with the English language and with their “poor previous educational environments.” Or else, why would my students return each semester to try to pass the CUNY-ACT? They want to pursue their dreams; their struggle represents the classic pursuit of success by way of the American Dream.

Pedagogical Implications

Before the arrival of the CUNY-ACT test, the Kingsborough English Department had implemented remedial courses based on portfolio assessment. The courses were designed to emphasize revision. Instructors and students worked toward the exit requirement of a portfolio that included multiple drafts, reading-based in-class exams, and a departmental reading-based final exam. Taken together, these assignments formed the basis for a reasonable and fair assessment of what the students were learning during the semester.

In the middle and at the end of each semester, instructors were able to discuss among themselves the progress of individual students according to departmental criteria for passing. Instructors could also work with their students on specific areas of writing that needed attention. Both instructors and students understood the requirements. The portfolios challenged students to produce personally meaningful writing that had to meet the demands of academic discourse, providing the basis for what Moss (1994b) called “more complex and authentic writing assessments” (p. 109). Students grew as writers through frequent student-teacher interaction; revision of drafts; classroom discussion of ideas. Assessment of this writing was then able, in Moss’ words, to “provide students with the opportunity to explore more of their own purposes, to rethink and revise their work over extended periods of time, drawing on existing resources and responses from readers” (p. 109). This system was well designed to help students develop their writing.

However, in practice, the system was not perfect. For one thing, “extended periods of time” were not always present. The remedial courses were linked, but only one semester long. As Sternglass noted, students lead very complicated lives, and as such need patience, time, and attention to learn to write confidently in order to progress. It is difficult for students to develop their writing in one semester, especially if they work full time, or meet other outside responsibilities. Sternglass believes that “[i]nstructors need to recognize that students must develop proficiency incrementally in both areas [content and form] simultaneously and that both

areas need to be stressed in subsequent courses over subsequent semesters” (p. 115). Although these remedial courses were specifically intended to help students “develop proficiency incrementally,” often they failed or were pushed along too quickly before they were ready by well-meaning instructors who felt that that they could work out their problems in freshman English. This was especially true for ELLs who were mainstreamed too quickly into non-ESL remedial courses. Before these problems could be addressed, the CUNY-ACT writing test arrived. Instead of providing students with “more complex and authentic writing assessments,” the CUNY-ACT writing test prescribed more simplified, standardized exit results that destroyed the gradual progress toward a better portfolio system.

The CUNY-ACT replaced the portfolios as the sole exit requirement for remedial students to pass into the first semester freshman composition course. In doing so, it immediately cut short the painstaking work students did in preparing the portfolios, and, in fact, rendered them meaningless. An added difficulty was that the ACT took effect in the middle of the Fall 2000 semester when the portfolio exit requirement had already been established and explained to the students. The English Department had to begin revising and explaining the change in exit requirement very quickly to confused faculty as well as to confused students, some of whom began to wonder if the department was developing a conspiracy to purposely fail them. The massive change in exit policy served to erode the confidence of the students and the students’ trust in their instructors whom they saw as part of the conspiracy.

As I explain more fully later, the ACT test was initially designed and implemented at the behest of local New York City politicians who charged a special CUNY Task Force Committee with creating an exit test for remedial students because some students at a Bronx CUNY campus, they claimed, were graduating with poor writing skills. Instead of conferring with the English Department at this campus to review its remedial course sequence, these politicians unilaterally and much too precipitously urged the task force to develop a strict one-shot testing program for all CUNY remedial students at all CUNY campuses that promised more immediate results, instead of proposing a solution that might have, in Moss’ (1994b) words, reflected a compromise “between competing criteria” for assessment (p. 110). The result was the creation of the CUNY-ACT writing test, a more cost-efficient, standardized procedure that emphasized reliability. Moss wrote: “Recommendations for enhancing reliability, without increasing the number of tasks or readers beyond cost-efficient levels, have typically involved increasing the degree of standardization in one or more aspects of assessment” (p. 112).

The ACT is an impromptu timed test for which students cannot prepare. They are given a prompt and asked to write a letter in 60 minutes to an administrator in response to it, the perception being that, in addition to or in spite of what they are currently learning, students should be able to write a simple letter in an hour. The ACT test was meant to cure or kill the patient (student) in 60 short minutes. As Brossell (1996) stated about the “state leaders” in Florida, such tests as the ACT are meant to provide “accountability” and ensure “decent standards of academic competence,” a sure cure to what such officials construe as a vote of no confidence in the academic community to assess their own students’ writing (p. 25). The “cure”

instead made matters exponentially worse. At Kingsborough, complications arose immediately.

For one thing, instructors learned that students in one particular remedial course, English 93, could take the ACT test *before* handing in their portfolios. The portfolios were evaluated separately from the ACT test. If the students passed the ACT test but failed the portfolios, they could progress to regular freshman composition. When students learned that they could do this, they resorted to haphazard work on the writing for their portfolios. Some students did no work on the writing and didn't submit the portfolios, thinking that they could get out of the course and into freshman English by passing the ACT. This scenario played out in some cases. Both students and instructors realized that the focus of the remedial courses tilted towards preparation for the ACT test. From 2000 on, instructors felt pressured to begin helping their students to study for the test, cutting short preparation for the portfolios. Students began to apply direct pressure because they knew that the ACT test was more valued than the portfolios. They wanted practice. As Moss (1994b) stated, "when assessments are visible and have consequences for individuals or programs, they alter educational practice, sending an unequivocal message to teachers and students about what is important to teach and learn" (p. 112).

This concern for the ACT test increased because of its high stakes in determining the future of the students who took the test. Moss (1994b) cited a paper presented by the National Commission on Testing and Public Policy which concluded that, "when the stakes are high—when schools ratings and budgets for teachers' salaries depend on test scores—efforts to improve performance on a particular assessment seem to drive out most other educational concerns . . . to progressively restrict curricular attention to the objectives that are tested and even the particular item forms that will appear on the test" (cited in Moss, 1994b, p. 113). Some instructors approached the problem by simply ignoring preparation for the test. Other instructors tried to regain their students' confidence and trust by searching for ways to prepare them for the test, without knowing the actual prompts the students would be given. There was no actual content for the students to study, which they could directly apply to the writing of the test. Instructors and tutors reverted to desperate, but questionable, prescriptive advice for how to respond to a particular prompt: "Don't write a counterview; the readers don't look for one"; "readers want to see that you have three distinct reasons in the body of your letter"; "don't copy the prompt word for word in your introduction." Most of this advice, in some cases contradictory, originated from hearsay about what CUNY-ACT readers looked for without anyone actually knowing how the tests were assessed, confusing the students and making them more anxious. The overall level of work declined. As Sternglass noted, "[u]nder timed, impromptu testing conditions, it is not possible to demand the level of writing competence that could be expected from college graduates" (p. 151). The CUNY-ACT test took away this challenge.

The readers who score the tests do not know the students, their backgrounds, nor the history of their writing development. They see only a name on a test. They certainly do not know whether the students have passed their portfolios or not. They are not trained to look into these areas. The ACT is governed from beginning to end by people trained as psychometricians who, in turn, train teachers to read

the tests and score them holistically. Remediation is then reduced to standardization, where, as Huot and Williamson (1997) explained, “individual achievement is decontextualized and standardized, so that tests can draw generalized inferences about individual performances on a particular test. . . . The emphasis is on the technical rigor of testing procedures and statistical operations and explanations rather than the complexity of student performance and judgments about that performance” (p. 46). Because of the test’s high stakes, preparation for it prevailed over preparation for the portfolios despite the pleas of well-meaning Kingsborough English Department administrators urging instructors not to “teach to the test.” However, this practice was no different from what occurred in Florida, where, as Brossell found, “in community college after community college, teaching to the test has become the accepted norm in composition classes” (p. 28).

Nevertheless, the messages that filtered down to Kingsborough’s instructors and students became mixed, at best. On the one hand, students needed to keep working on their essay writing to fulfill the requirements for the portfolios, a well-established requirement; on the other hand, they needed to learn how to pass the ACT test. In this case, standardization won. Huot and Williamson explained the system this way: “Large-scale, high-stakes testing requires standardization and tends to reduce the curriculum to what can be measured. At best, test scores obtained under these conditions are a very poor indicator of the range of learning fostered by a school curriculum” (p. 46). If students complete a remedial course sequence where the emphasis is on preparation for the ACT, and if they pass the test, they then enter freshman composition without the preparation necessary for work in the course. They spend too much time preparing for the ACT test, without receiving the extended help that might have helped them prepare for the next course sequence. But if instructors try to help their students with portfolio preparation, the students work with little motivation, wondering how the drafting and revising will help them prepare for the ACT test. The test becomes a prescription for failure in a once viable program that was based on sound pedagogical theory and practice.¹

Kingsborough’s program derived from those who believed in the capability of remedial students to perform well in college-level work, with the right mix of motivation and support. Gleason (2000), who tried unsuccessfully, in the face of heavy political opposition, to establish a policy of enrolling remedial students into regular credit-bearing college courses at CCNY (Sternglass’ college), discovered that the students whom she did manage to shepherd through, “were passing the core courses at a rate that was even higher than the rate for our pilot course students who had placed into English 110” (p. 568). She evidently believed enough in her students to predict their success with much greater accuracy, without the aid of the CUNY test that preceded the CUNY-ACT: The CUNY-WAT Writing Assessment Test (WAT): “Our college’s policy of placing students who fail the CUNY Writing Assessment Test into a (reduced credit or no credit) remedial writing course and barring them from required core curriculum courses is underwritten by the assumption that the writing test predicts students success in college courses” (Gleason, p. 569). She successfully challenged that premise, but the college, operating under the assumption of “deficiency” commonly associated with

“remediation,” ignored her research results and chose to rely on the test (see the full article for a discussion of the circumstances surrounding her efforts).

Gleason was trying to dispel the notion that students who entered CCNY via open admissions were underprepared for college-level work, which, for many in the profession meant that they were cognitively impaired (see Rose, 2001, on the relationship of school performance to cognition). At the beginning of open admissions at CUNY in the late 1960s and early 1970s, pioneering teachers such as Shaughnessy (1977) saw clearly that open admissions students were not cognitively deficient, but, rather, were not taught how to write in their secondary schooling. According to Shaughnessy, “BW [basic writing] students write the way they do, not because they are slow or non-verbal, indifferent to or incapable of academic excellence, but because they are beginners and must, like all beginners, learn by making mistakes” (p. 5). She saw that these students “were in college for one reason: that their lives might be better than their parents’, that the lives of their children might be better than theirs so far had been” despite having “grown up in one of New York’s ethnic or racial enclaves” (p. 3). Many had spoken other languages or dialects at home and never successfully reconciled the worlds of home and school, a fact that by now had worked its way deep into their feelings about school and about themselves as students” (p. 3). In short, Shaughnessy realized how open admissions students were different in terms of their home and school environments. Shaughnessy’s view then was that, given the proper guidance and support, these students could begin to overcome these deficiencies and gain confidence in themselves as learners. Equality of opportunity was important then. (For more on Shaughnessy, see Maher, 1997, especially pp. 249-310 for a compilation of Shaughnessy’s writings.)

The consequences of quick-fix standardized testing are serious. As Scharton (1996) noted, a more draconian view of testing has become prevalent, based on a social Darwinist theory of natural law (p. 70). According to Hull and Rose (1991), remedial students are socially constructed as “cognitively defective and in need of ‘remedy’” (p. 299), as implicitly impaired beyond anyone’s help. Scharton’s position reinforces a belief common in today’s recidivist philosophy that society is inherently unequal, that society always has “a top, middle, and bottom” (p. 70); therefore, the view goes, “it’s no one’s fault if some people fail” (p. 70). I would go farther and say that this view presumes that “some people” *must* fail in order to sustain this prevailing view. Equality of opportunity is now deemed not so important, although college administrators pay lip service to this possibility. They would not be unhappy to see remedial students leave. Quickly.

Gleason and Sternglass have both emphasized that students who need patient help in developing literacy skills need more instruction over a longer period of time especially if the college commits itself to these students in the first place. However, according to Scharton, this social Darwinist view has influenced English teachers as well. He reported hearing English teachers praise timed written essay tests “to catch those illiterate students their lazy colleagues were letting through the system” (p. 72). Scharton found that “the students found by the test to write less well than the norm were ones whom the academic system in general and the writing program [at his school] in particular had already marginalized, even terrorized” (p. 72).

In trying to help students pass the ACT test, I try my best to focus in helping students to pass the ACT test. They are caught in a terrible political situation. I am sympathetic. I try to remain positive. I want my students to succeed. I do not look at anything else: What happens after they pass; how they feel about writing afterwards. I know that their struggle is not over. Once they pass the test, I feel I have done my job. But the price to pay for this focus is high. As Moss (1994b) contended, “Although high stakes testing programs frequently result in improved test scores, such improvement does not necessarily imply a rise in quality of education or a better educated student population” (p. 114).

The CUNY-ACT

The students are presented with two prompts, each involving grant money specially targeted to improve some aspect of a (a) particular community and (b) a school. Each prompt presents two proposals, only one of which can be chosen because the grant money is limited. The grant is to help benefit the quality of life for a specific part of the community—or school—population. The students must first decide which prompt to respond to, the community-based or school-based prompt, and then pick the proposal that seems to them most beneficial for this particular population and write a letter to the appropriate person (a community board member or town mayor; or, if school-based, to a dean or principal) to convince him or her why the money should be spent on this particular proposal. The students must plan, organize, and complete their letters within the prescribed time. Here are two examples of CUNY-ACT prompts, one a school-based prompt, and the other, a community-based prompt:

Example #1

Your college has received a large donation from a recent graduate who has specified that the money be used to enhance the quality of life for students. Two proposals for using the money have been made. One proposal is to expand the campus bookstore to include music CDs, gifts, and a snack bar. The other proposal is to add more computers to the computer lab. There is enough money to fund only one of these proposals. Write a letter to the dean of the college in which you argue for expanding the bookstore or for adding more computers to the computer lab, explaining how your choice will enhance the quality of life for students. Begin your letter: Dear College Dean:

Example #2

A local community board wants to start a program that will increase cultural cooperation and awareness in the neighborhood. Two proposals have been made. One proposal is to sponsor a multicultural street fair including cultural music, street performers, and vendors selling ethnic foods and goods. Another proposal is to start a program called “Meet Your Neighbor,” in which community residents from different backgrounds would talk to people at the local libraries to share information about their cultural practices and traditions. There is enough money to fund only one of these proposals. Write a letter to the local community board in which you argue either for sponsoring the multicultural street fair, or for beginning “Meet Your

Neighbor” talks at the local libraries, explaining how your choice will help to increase cultural cooperation and awareness in the neighborhood. Begin your letter: Dear Members of the Community Board:

When the tests are completed, they are collected and shipped off to another CUNY college where they are all read within a few days’ time by readers from CUNY campuses, teachers who sign up to be readers, and are then “normed” during large all-day sessions where they read and comment on CUNY-ACT sample tests (a process that is meant to ensure interrater reliability). These teachers are paid a fee by a company contracted by CUNY Headquarters to administer the test. The tests are read through twice, holistically, and assigned a numerical grade ranging from 1 to 6. (The most common grades assigned are 3 and 4.) A second reader of a test does not know the score of the first reader. Then the scores are added up. A two-reader sum score of 7 is passing (a reader who scores one test a “3” and the second one who scores it a “4,” or *vice versa*). There is no resolution if there is a split score of 3 or 4. Third readings have been eliminated. However, there can be no score of 7 if the split is, say, 2 and 5. More than one point of disagreement just does not exist. Hence, most readers are urged to seek a middle ground in their judgments. The rubric controls the assessment, which, as Broad (2003) stated, is exactly how the system is supposed to work, “since their [rubrics] chief purpose is to constrain the range of criteria and thereby boost interrater agreement” (p. 25). Therefore, a well-written test usually gets a 4, and a less-well-written test, a 3. Sometimes a test will receive two 5s, totaling 10, although this instance is rare. Generally, tests are scored as either a 6 or a 7. If two readers score one test a 3, a total of 6, then the test fails. After all the tests are read, the scores are tabulated by computer and sent to the appropriate CUNY college English Department. The tests are then stored away into a vault. They are not sent back to the college from whence they came. (See appendix for scoring criteria.)

If a student fails the exam, he or she can appeal, but only if the student receives a score of 6, nothing lower. The appeal process involves having a chief reader (an English Department member at my college) find the exam, read it over, and then fill out a form explaining to the student why he or she received the score that was assigned. The student meets with the chief reader, but cannot see the exam. This rule presents a problem when I work with students in English W who have failed the test. I can work with the students in my course and try to help them, but I don’t know what they did wrong on the test because the results are simply noted by number. I do not have access to the tests that the students have failed. This stipulation means that when students enroll into an English W course, they—and I—do not know why or how they failed. We must then approach work in further preparation for passing, from scratch.

English W—The CUNY-ACT Prep Course

I approach English W as the test-prep course that it is intended to be. I work specifically to help prepare students for the test. I tell them that the test is based on a particular form, that it is similar to a business letter written to an admin-

istrator, which must have standard written English that presents an argument for a particular proposal in terms of details, facts, numbers, suggestions, ideas that an administrator wants to see if he or she is going to part with a large amount of money to fund a particular proposal. I produce for them a template of what a letter should look like. We read over models that I have accumulated and have written up myself to point out different ways of approaching a prompt. We review their contents so that the students have a clear idea what distinguishes a passing test from a failing test. I share with them the scoring criteria that are given to me as well as the rubrics. I am also given sample tests that correspond to a particular rubric, which I also share with them.

After discussing sample letters and their evaluations, I present to them other samples that are not evaluated. I then form the students into groups to evaluate the samples, asking them to note the “strengths” and “weaknesses” of each sample. Eventually, I give them practice tests. I pick and choose from many practice prompts that are stored in the English Skills Lab and in a booklet designed to guide students in preparing for the exam. I often ask them to indicate whether they would like me to type them up for general class review the next day. I type them up, photocopy them, and then distribute them to the whole class. The students do the same with these that they did with the models: Evaluate the typed up exams, noting their “strengths” and “weaknesses.”

There are advantages to this procedure. The students get to talk about their writing, about writing in general, and they begin to coalesce as a group. They learn what works best in the time they have to write the exam, and put into practice some planning tactics I emphasize: “Brainstorming,” “clustering,” or outlining. They bond together as a group in finding themselves locked into a common perplexing situation: They are all nervously waiting to take the real CUNY-ACT. As the time nears for the real exam, they become nervous and anxious, but by then, they express their fears and support one another.

The results of these efforts are disappointing, sometimes predictable, sometimes not. Overall, I have found that about half of all the students I teach in English W pass. For example, in a short, 6-week Summer 2005 English W course, euphemistically called an “immersion” course, only 3 of 12 students passed the CUNY-ACT. In the 6-week Winter 2004 “immersion” course, I had 24 students in my English W, 18 of whom were non-native English speakers and writers. Fifteen students passed the CUNY-ACT. In the Summer 2004 English W course, I had 22 students, 15 of whom were non-native English speakers and writers. Nine students passed the CUNY-ACT.

The results are monitored by the Office of Institutional Research at Kingsborough (2007) which tallied the scores. This office found that the pass rates of students who took the ACT as an exit exam from Fall 2000: Fall 2000: 37.4; Fall 2001: 40.4; Fall 2002: 41.4; Fall 2003: 43.7; Fall 2004: 47.2. The pass rates for all CUNY community colleges hover around 50%. In real terms, for example, this means that of 1,132 students tested in Fall 2003, only 495 passed (43 %). The Office of Institutional Research does not explain the rationale for the range of statistics, but the scores seem to reflect what Scharton (1996) called “a normal distribution of scores” consonant with the Bell Curve (see Scharton’s discussion on pp. 69-70, to

which I alluded earlier). The fact that more than half the students who took the test failed would seem normal according to this view. A Bell Curve explanation would seem appropriate in evaluating the results in light of the fact that so many students continue to fail year after year. It does not seem so unusual then that most of my students in English W have taken the test at least twice. They were included in the majority of students who did not pass, one semester after another. The low pass rate is particularly disturbing. Not once was there a pass rate of more than 50%.

Nevertheless, I am gratified to know that some of my students pass but perplexed as to why others did not. I never know what they actually did because the tests are stored away. When I see the students after the test, many say they felt so nervous by the autocratic conduct of the people who administered the test, that their anxieties increased rather than diminished, and they simply froze up.

The Origin of the CUNY-ACT Test at CUNY

The CUNY-ACT tests were created during a tremendous New York City political storm that roared through in 1997 and 1998 when several things happened at once. First, in late Winter 1997, New York's governor, George Pataki, along with the CUNY Board of Trustees (10 of whom had been recently appointed by Governor Pataki and by then New York City mayor, Rudolph Giuliani) became concerned that many CUNY students were performing at high school-level achievement (Arenson, 1997b). Governor Pataki and Mayor Giuliani were joined in their criticisms by Herman Badillo, Mayor Giuliani's education advisor, recently named vice chairman of the CUNY Board. All three men wanted something done about CUNY's remediation programs.

Their language was not temperate. As Arenson (1997b) reported, Herman Badillo stated: "City University is supposed to be a university of opportunity to move into the labor force, not a remedial high school" (p. B3). Giuliani used stronger language. Reacting to the fact that only 5% of CUNY's community college students graduate in 2 years and that only 19% in 5 years (as reported in Arenson), Mayor Giuliani called the statistic, "absolutely pathetic," while Governor Pataki is reported to have said that the record is "dismal" (Arenson, 1997b). The self-righteous indignation expressed in these statements does not take into account why many of these students take so long to graduate: According to Sternglass' ethnographic study of nine undergraduates, "the amount of time needed to complete degree requirements continues to grow as the economic support for poor students declines, necessitating longer time periods to complete academic programs" (p. 296). Realistically speaking, if colleges are going to commit themselves to working with students who need help, then they must be prepared to patiently spend the time and money to do so. But this possibility seems dim in the current state of educational funding that has been cut back, not increased.

Badillo had become especially upset at the fact that some community colleges were not administering the CUNY-wide CUNY WAT, which preceded the ACT. The WAT was a standardized test that differed from the ACT in that it presented students with a prompt about a controversial subject (e.g., abortion or the death penalty), and then asked students to "agree or disagree" based on their knowledge

and experience. Badillo said he believed in standardized tests as effective measures of writing skill even though the WAT was not meant to be an exit exam. As John Mayher, professor of English education at New York University, stated, "It [the WAT] was not designed as an exit criteria but as a diagnostic test" (Gonzalez, 1997, p. 3).

At this point, Badillo and the Board became much more assertive in trying to implement their own ideas of exit criteria for the CUNY colleges. They criticized the chancellor of CUNY, Dr. W. Ann Reynolds, for the problems at CUNY, ultimately forcing her to resign. Their concern was that CUNY professors were not doing their jobs and that they were inflating grades (see Arenson, 1997a). Throughout Spring 1997, a climate of hostility prevailed within the CUNY system, precipitated by the intense, aggressive involvement of Badillo and the chair of the trustees, Dr. Anne Paolucci. The level of acrimony escalated, prompting a lead editorial in the *New York Times* that read, "Intramural Warfare at CUNY" (July 7, 1997), which essentially tried to sort out the situation without delving into the educational issues at stake.

The second crisis occurred a year later when the Board demanded that all remedial programs in CUNY's 4-year schools be phased out entirely (Arenson, 1998; Healy & Schmidt, 1998). According to Badillo, the move meant that 4-year CUNY colleges would gain immediate recognition as institutions with high standards. "We're moving back to having academic standards that preserve the value of a CUNY diploma," said Badillo (cited in Healy & Schmidt, 1998, p. A21). This move also meant that the community colleges would inherit the work of remedial education, a trend followed nationally, making public education less accessible for minority and disadvantaged students from the New York City area, a fact publicly acknowledged by Paolucci: "We are cleaning out the 4-year colleges and putting remediation where it belongs" (cited in Herbert, 1998, p. A29). What is evident in this quote is the use of the term *cleaning out*, which, unfortunately, resonates with a much more pernicious term, "ethnic cleansing," implying that the students need to be "cleansed." Such a view corresponds with the belief I noted earlier that remedial students are socially constructed as "lesser in character and fundamental ability" that suggest a "flawed character" (Hull & Rose, 1991, p. 311). Elsewhere, Rose (1985) noted that remedial students are viewed as being diseased, having a "mental defect" (p. 349) instead of the term *deficiency* being associated with poor secondary schooling, discussed by Shaughnessy. (For an interesting historical perspective about composition's "narrative of lack" in constructing a view of teaching and of students to "improve," with regard to their literacy, see Slevin, 2001.)

These two policy changes regarding remedial education during Spring 1999 led to the publication of a report by the Mayor's Task Force on CUNY, a committee set up by the mayor to issue findings on remediation in CUNY. The report was ominously entitled, "The City University of New York: An Institution Adrift." In its opening pages, the committee reports its commitment to CUNY as a "model of excellence" as "the preeminent urban public university in the world" (Schmidt, Badillo, & Brady, 1999, p. 5). How should CUNY go about achieving this goal? By establishing rigorous "clear" standards that according to the committee are not being met basically because of a remediation system that is "flawed." The language

is specific: “Thirty years after the implementation of open admissions, CUNY has not yet established valid and reliable remediation tests. It does not carefully diagnose students’ remedial needs. It does not measure objectively what students have actually accomplished in remediation, nor has it promulgated systematic and valid standards to determine what students may exit remediation” (p. 7). The committee wants CUNY to establish an assessment program that “is consistent with modern assessment science” and that accounts for “objective standards” (p. 7).

Although the report stipulates the need for “valid and reliable remediation tests,” the phrase itself tends to conflate the distinction between validity and reliability. The call for “objective standards” places the emphasis on reliability with strict interreader reliability, in assessing the ACT results. Moss (1994a) has specifically stated that educational testing companies such as the one that has established the ACT at Kingsborough, pay more attention to “continued reliance on reliability, defined as quantification of consistency among independent observations, requir[ing] a significant level of standardization” (p. 6). As I noted previously, the assessment of the ACT tests depends exclusively on “normed” readers who are trained to score the tests quickly, efficiently, objectively, seeking consensus in scoring, without once inquiring into the identity or history of the students whose tests they are evaluating. Readers do not see the students as individuals with varying degrees of potential or capability. Therefore, committee members apparently did not heed Moss’ (1994b) advice that “policymakers have been too quick to implement assessment systems without adequate attention to the potential and actual consequences of their actions” (p. 124). She further warned of the “unintended consequences” that such assessment designs may cause. The call for “objective standards” tends to glide over the issue of unintended consequences, in its urgent appeal for efficiency (i.e., reliability vs. the more important consideration of validity). One such unintended consequence is the “undermining of progress in areas not addressed [in the outcome assessment]” (p. 123).

Validity, as posited by Cronbach (1988) and Messick (1988) and adopted in the joint publication on professional guidelines and standards for testing by the American Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999), presents a much different construction of testing and test results. According to Cronbach, validity is more of a debate, an argument, that does not correspond to the drive for “objective” results. Huot (2002) supported this view. Quoting Messick, he stated that “[v]alidity . . . has evolved from a simple correlation to ‘an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessments’” (p. 49). Validity is not meant to be objective, is not at all similar to seeking the same results that are posited in quantifying test results that rely upon “interreader reliability.” Huot asserted that validity is, rather, “an ongoing process of critical reflection” (p. 51) and differs from reliability’s drive for efficiency, in that it seeks “to represent complexity and variety inherent in written communication” (p. 52). Cronbach’s view is that “[v]alidation speaks to a diverse and potentially critical audience; therefore, *the argument must link concepts, evidence, social and personal consequences and values*” (p. 4; italics in original). My

point is that not once, in its report, has the Schmidt Committee acknowledged or cited recent findings in the area of “modern assessment science” although its assertions imply familiarity with and understanding of the latest research in educational measurement. Had the committee members consulted the current and relevant literature on professional uses of educational tests, they would have understood that any proposals that purport to rely on scientific measurement are made essentially in a political context: “[T]he practical use of measurements for decision making and action is or ought to be *applied* science, recognizing that applied science always occurs in a political context” (Messick, 1988, p. 43). Messick stated categorically that validity is not to be thought of in a strictly scientific context, but said, “the justification and defense of measurement and its validity is and may always be a rhetorical art” (p. 43; italics in original). Nowhere in its report does the Schmidt Committee members acknowledge or reference the existence of professional guidelines and standards for the use of educational tests.

Yet, with a sense of misguided confidence, the committee members assert that their idea of effective assessment tests will help remedial students to either graduate, to be pushed into vocational education courses (Schmidt et al.), or to simply drop out altogether. In one section of the report about assessment testing, the committee recommends the following: “CUNY should immediately discontinue use of the WAT and, as a stopgap measure, replace it with either the College Board’s Accuplacer or ACT’s Compass writing assessment test. . . . CUNY must implement specialized assessment for ESL students. CUNY must back its remediation exit standards with valid and reliable post-tests that can determine whether students are in fact ready for college-level work” (p. 39). The assertiveness of its appeal camouflages the deleterious effect of its proposal, failing to recognize that validity is about the decision(s) made on behalf of a test.

There are no warnings in the report about social and personal consequences for the students whom the committee members so eagerly want to test, as made quite evident in Paolucci’s statement about “cleaning out the 4-year colleges” and relegating those students who need more patient help and attention, to 2-year colleges. Nowhere in their report have they asked such crucial questions that any assessment specialist would want to know in designing a test on which valid decisions can be made. Shepard (1993) contended that validity is an important part of designing any test: “My proposal . . . is that validity evaluations be organized in response to the question ‘What does the testing practice claim to do?’ Additional questions are implied: What are the arguments for and against the intended aims of the test? And What does the test do in the system other than what it claims, for good or bad? All these issues should be sorted through at once, with consequences as equal contenders alongside domain representativeness as candidates for what *must* be assessed in order to defend test use” (pp. 429-430). Nowhere in the Schmidt Report are these considerations noted. The Schmidt report seems content to simply assert the need for “clear standards of readiness for entry into college-level work and corresponding remedial exit standards” (p. 99) by instituting the ACT as a “placement, diagnostic, and post-test instrument[s]” (p. 100) without exactly understanding what that entails, not even acknowledging its rationale about why a placement test is to be also used as an exit test.

Additionally, the committee members have not considered the differences in pedagogical systems established within the individual CUNY colleges. The language of the Schmidt Report presumes a one-size-fits-all concept. The use of “objective” tests, with “clear standards” that are “systematic” with “accepted scientific standards of reliability, validity, and fairness” presumes that all of the assessment tests, all of the policies at each individual college, will be the same (Schmidt et al., p. 26). But there seems to be no mention that each CUNY campus has a diverse, different student population, and that each English Department has set up unique programs suited for its student population. Therefore, once again, there is evidence that the committee exhibits no understanding of the “social and personal consequences” that such a testing policy would have on these students, many of whom, as ESL students, have such difficulty writing a cogent, persuasive essay in 1 hour because of their severe limitations in the English language. It is simply unjust and unfair. Cronbach is relevant here: “*Tests that impinge on the rights and life chances of individuals are inherently disputable*” (p. 6; italics in original). The intent seems to be to punish and shut out ESL students who have not learned quickly enough, but who, given time and patience, will learn to overcome their difficulties, even though there is mention made in the report about ESL students: “CUNY must implement specialized assessment for ESL students” (Schmidt et al., p. 39; I elaborate later on this issue.) Although the language in the report purports to sound fair, the intent is apparent: To sort out those students who do not meet the “clear standards” haphazardly conceived by the committee (see Fraizer, 2003; Scharton, 1996, pp. 70-74, for issues related to political meddling in educational theory and practice).

The report ends by calling for CUNY to commit itself to “a unified, coherent, integrated public university system, even though this difficult undertaking is unprecedented in its history” (Schmidt et al., p. 91). In effect, the committee wants, above all, uniformity, as well as excellence. The committee seems to see, to literally perceive, all of the CUNY colleges as the same, with a student body that must meet similar standards and exit criteria.

Pedagogical Consequences

This situation at Kingsborough illustrates the prevailing view that writing assessment is indeed, as White (1996) noted, “a site of contention.” When Schmidt et al. produced their 1999 report, they recommended that assessment be placed in the hands of test professionals with their own set of values and guidelines about testing. These goals clashed with the goals and values of writing instruction at Kingsborough, which had already established a systematic set of requirements designed to meet the needs of its individual student population, based on current composition theory and practice.

College English departments like the one at Kingsborough design pedagogies and assessment based on recent research in the field of composition studies, following guidelines according to the context of their pedagogies, a formal statement of which is readily accessible in the CCCC Committee on Assessment (1995) made more than 10 years ago (see also Hamp-Lyons 1995). Without knowledge of these guidelines, administrators like those on the Schmidt Committee risk constructing

“mismatch[ed]” assessment constructs, between those based on current theories espoused by compositionists and those “measured in traditional formats for assessments” (Camp, 1996, p. 136). This mismatch is what Camp, citing Messick (1989) called “construct-irrelevant variance,” where the assessment constructs inadequately represent writing as those in composition know it and where even multiple-choice tests would seem logical, but not at all valid in relation to a particular department’s pedagogy (Camp, p. 137).

The danger of such constructs is most serious for ESL students who make up such a large percentage of the student population at Kingsborough. Their needs are not being met in preparing for and in taking the ACT because, as Camp asserted, of “the possibly deleterious effects of conventional writing assessment formats [such as the ACT] on students outside the mainstream of academic culture” (i.e., ESL students). Most ESL students fail because the test does not take into account their need for a continuing, comprehensive, and appropriate program of instruction and evaluation that will challenge them to learn to write in English. The ACT test creates more obstacles for them, rather than presenting occasions for them to become better writers. They cannot learn from failing the test, which is stored away in a vault. If they pass Kingsborough’s course requirement for its remedial course, but fail the ACT, all they can do is to take English W and hope to practice enough to pass the test when they must retake it under the exact same conditions as everyone else; as such, they are caught in a maze in this high-stakes assessment procedure, further damaging their ability to learn to write in English, since learning English is reduced to passing a single test. This continuation of failure further erodes their level of confidence but not their determination to pass. Messick (1989) warned that the “side effects” of testing should be closely monitored, that “relevance and utility as well as appropriateness of test use depend, or should depend, on score meaning (p. 8). To act otherwise is not just dubious but dangerous. Using test scores that ‘work’ in practice without some understanding of what they mean is like using a drug that works without knowing its properties and reactions” (p. 8). Therefore, the need is great for a more balanced, contextualized assessment procedure that takes into consideration the social consequences for the student population being tested, especially ESL students.

Assessment should optimally grow out of pedagogy, not out of the need for accountability. However, professional testing services look at nothing else but assessment, and as such, they value, above all, accountability. As Elliot (2005) observed, “[t]he drive for accountability leads to an efficiently designed assessment that, in turn, leads to a construct of literacy that is reified from the design—a solipsistic nightmare” (p. 352). The “design” creates the context of assessment, whereas, the reverse should be true: The “design” of the assessment should follow the requirements of locally designed curricula. What has happened in the case of the CUNY-ACT at Kingsborough was that the cart came before the horse. Assessment and its need for efficiency drove the curriculum. Williamson (1994) has explained, “the historical emphasis on efficiency has continued to favor reliability over validity” where reliability “has had a highly deleterious effect on pedagogy, since most researchers or teachers of writing would see such a presumed definition of writing as extremely limited” (p. 165).

The CUNY central administration based its presumptions on misguided perceptions of remedial education and of validation in assessment. The result is an unethical abrogation of a college English Department's right to design its own assessment procedures. As Williamson and Huot (2000) stated in their discussion of ethics in assessment:

Ethical uses of writing assessment dictate that we focus outward toward the students we test and the ramifications of this testing rather than inward toward the technical and statistical properties of the test itself ... At their worst, the tendency of assessment technology to become a form of gatekeeping can work to socially engineer and perpetuate inequalities in American society as they replicate the inequities of an imperfect system. (p. 20).

CUNY Central should be supporting, not undermining, Kingsborough's English Department to improve its remedial course system.

Students, instructors, and English departments need to know that the work they are doing is constructive, valid, and worthwhile. Checks and balances are necessary in ensuring that the work is also justifiably challenging and productive. CUNY Central headquarters can help the Kingsborough English Department in this regard by inquiring into what has proven successful and what needs improvement. Should more time be given for remediation? Does there need to be greater oversight in assessing departmental final exams? Are the criteria for passing reasonable, fair, and challenging? To address questions such as these, I suggest the formation of a standing committee composed of equal numbers of CUNY Central administrators, English Department faculty, and students to review random samples of the portfolios and/or the departmental final exams of students who passed and who failed the exams at the end of each semester. Were the exams evaluated fairly? How did they correspond to the writing in the portfolios? Assessment might then be a cooperative effort.

The bargain would be that there would be no more CUNY-ACT or anything resembling a standardized timed test. But the department would have to justify the worth and validity of its portfolio system to CUNY Central headquarters. There would be a redesigned English W course in which instructors could work with students on the aspects of their writing that need attention, from having failed the departmental writing test, but they would have to pass the departmental final after perhaps two years time. Students would be able to take the tests with them to this class for review so that they would know how to improve when they are retested. English W would then become a constructive workshop and legitimate extension of the work done in the previous remedial courses. The hope is that this redesigned method of assessment would help CUNY Central headquarters validate the work done in Kingsborough's remedial courses while at the same time overseeing it.

Whether CUNY Central headquarters would agree to negotiate with the English department in this manner is an open question. However, the students at Kingsborough Community College will not be well served unless the English Department faculty members are granted the right to help design exit requirements and course work appropriate to their student population, in conjunction with CUNY Central. Haswell and Wyche (2001) made quite clear that "writing teachers

should be leery of assessment tools made by others, that they should, and can, make their own” (p. 14). The best-case scenario would be that CUNY headquarters retain oversight in the progress of the CUNY students, but trust the individual English Departments to create their own placement and exit requirements. By the same token, departments such as Kingsborough’s need to assume more confidence and assertiveness in negotiating in good faith in trying to implement their own assessment policies with mandated reviews to chart their effectiveness. Moss (1994b) remarked: “Teachers need to assume more responsibility for accounting for their own practice through collaborative inquiry and ongoing peer review, so that their voices are not overshadowed by externally imposed assessments; administrators need to provide them with the time and resources to do so” (p. 124). CUNY Headquarters needs to listen to individual CUNY English Departments. The goal then remains for each department to actively set up and campaign for its own placement and exit criteria in negotiation with CUNY headquarters, employing Huot’s (2002) “validity as argument” approach, and to do so with the courage of conviction based on years of experience and knowledge in instruction and in assessment (p. 56).

I am now beginning the Spring 2008 semester teaching another English W course. In this course, there are students from many different countries: Poland, Pakistan, The West Indies, Albania, the Ukraine, Russia, China, Egypt, Senegal, Puerto Rico, the Dominican Republic, Mexico. Only five are originally from the United States. In terms of having previously taken the ACT test, three students have failed the ACT four times, eight students have failed it two times, four students have failed it three times, one student has failed it five times, and one student has failed it six times. Last fall, one student in my English W had failed the test seven times. After taking my class, she passed it. She never gave up hope. I was very happy. My current students are optimistic that this will be the time that they pass as well. I will work hard to help them do so.

Appendix

Act Scoring Scale

Upper-range papers. These papers clearly engage the issue identified in the prompts and demonstrate superior skill in organizing, developing, and conveying in standard written English the writer’s ideas about the topic.

- 6 Exceptional. These papers take a position on the issue defined in the prompt and support that position with extensive elaboration. Organization is unified and coherent. While there may be a few errors in mechanics, usage, or sentence structure, outstanding command of the language is apparent.
- 5 Superior. These papers take a position on the issue defined in the prompt and support that Position [sic] with moderate elaboration. Organization is unified and coherent. While there may be a few errors in mechanics, usage, or sentence structure, command of the language is apparent.

Mid-range papers. Papers in the middle range demonstrate engagement with the issue identified in the prompt but do not demonstrate the evidence of writing skill that would mark them as outstanding.

- 4 **Competent.** These papers take a position on the issue defined in the prompt and support that position with some elaboration or explanation. Organization is generally clear. A competency with language is apparent, even though there may be some errors in mechanics, usage, or sentence structure.
- 3 **Adequate.** These papers take a position on the issue defined in the prompt and support that position, but with only a little elaboration or explanation. Organization is clear enough to follow without difficulty. A control of the language is apparent, even though there may be numerous errors in mechanics, usage, or sentence structure.

Lower-range papers. Papers in the lower range fail in some way to demonstrate proficiency in language use, clarity of organization, or engagement of the issue identified in the prompt.

- 2 **Weak.** Although these papers take a position on the issue defined in the prompt, they may show significant problems in one or more of several areas, making the writer's ideas often difficult to follow: support may be extremely minimal; organization may lack clear movement or connectedness; or there may be a pattern of errors in mechanics, usage, or sentence structure that significantly interferes with understanding the writer's ideas.
- 1 **Inadequate.** These papers show a failed attempt to engage the issue defined in the prompt, lack support, or the problems with organization or language are so severe as to make the writer's ideas very difficult to follow.

Note

1. The power of standardization to influence pedagogy is considerable. The English Department recently eliminated portfolios from English 93, but kept a reading-based final exam created by the Department, that was previously the last piece of writing to be included in the portfolios. This exam is now given a week before the CUNY-ACT exam, and preparation for it begins a week before that, when students are given the essay to read and discuss, on which they will be tested. This departmental exam is read and evaluated pass-fail by small "cohort" groups of English 93 teachers [with a "cohort" leader] who exchange and read one another's exams before the ACT exam. If students pass the final exam, they are given a grade for the course based on the essays they wrote during the semester plus the final exam. If they fail the Departmental exam and pass the ACT, they are given a grade of "F" ["repeat"],

but the “R” is meaningless: They proceed directly to freshman English. (Students can take the ACT if they take the final exam first.) The change in exam sequence was established to eliminate the possibility of students skipping the final exam a week after the ACT, thinking they will pass the ACT. Nonetheless, the ACT’s high stakes remain the same: Passing the ACT signifies entrance into freshman English.

Acknowledgments

I would like to thank Bill Condon, Norbert Elliot, John S. Mayher, and David Simms for their support and encouragement as I wrote and revised. I would also like to thank Pat Belanoff and Edward White for their very helpful suggestions in reviewing the manuscript, and special thanks to Brian Huot whose editorial expertise and guidance helped me to understand and to appreciate writing assessment in theory and practice. Finally, I would like to thank my students for their determination and perseverance. They inspired me to write about their quest to succeed in college.

References

- Arenson, K.W. (1997a, June 25). CUNY board’s leader says chancellor is under review and suggests she resign. *The New York Times*, p. B3.
- Arenson, K. W. (1997b, March 18). CUNY trustees question the value of remedial programs in math and writing. *The New York Times*, p. B3.
- Arenson, K. W. (1998, May 8). Critics fear role of politics in CUNY plan. *The New York Times*, pp. B1, B6.
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan: Utah State University Press.
- Brossell, G. (1996). Writing assessment in Florida: A reminiscence. In E. M. White, W. D. Lutz, & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, practices* (pp. 25-32). New York: Modern Language Association.
- Camp, R. (1996). New views of measurement and new models for writing assessment. In E. M. White, W. D. Lutz, & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, practices* (pp. 135-147). New York: Modern Language Association.
- CCCC Committee on Assessment. (1995). Writing assessment: A position statement. *College Composition and Communication*, 46(3), 430-437.
- Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer (Ed.), *Test validity* (pp. 3-19). Hillsdale, NJ: Erlbaum.
- Elliot, N. (2005). *On a scale: A social history of writing assessment in America*. New York: Peter Lang.
- Fraizer, D. (2003). The politics of high-stakes writing assessment in Massachusetts: Why inventing a better assessment model is not enough. *Journal of Writing Assessment*, 1(2), 105-121.
- Gleason, B. (2000). Evaluating writing programs in real time: The politics of remediation. *College Composition and Communication*, 51(4), 560-589.
- Gonzalez, J. (1997, May 29). ABCowards of the CUNY writing plan. *Daily News*, pp. 3, 67.
- Hamp-Lyons, L. (1995). Review: Uncovering possibilities for a constructivist paradigm for writing assessment. *College Composition and Communication*, 46(3), 446-455.
- Haswell, R. H., & Wyche, S. (2001). Authoring an exam: Adventuring into large-scale writing assessment. In R. H. Haswell (Ed.), *Beyond outcomes: assessment and instruction within a university writing program* (pp. 3-13). Westport, CT: Ablex.

- Healy, P., & Schmidt, P. (1998, July 19). In New York, a "standards revolution" or the gutting of public colleges? *The Chronicle of Higher Education*, pp. A21-23.
- Herbert, B. (1998, May 28). Cleansing CUNY. *The New York Times*, p. A29.
- Hull, G., & Rose, M. (1991). Remediation as social construct: perspectives from an analysis of classroom discourse. *College Composition and Communication*, 42(3), 299-330.
- Huot, B. (2002). *(Re)articulating writing assessment for teaching and learning*. Logan: Utah State University Press.
- Huot, B., & Williamson, M. (1997). Rethinking portfolios for evaluating writing. In K. B. Yancey & I. Weiser (Eds.), *Situating portfolios: Four perspectives* (pp. 43-57). Logan: Utah State University Press.
- Intramural Warfare at CUNY. (1997, July 7). *The New York Times*, p. A14.
- Maher, J. (1997). *Mina P. Shaughnessy: Her life and work*. Urbana, IL: National Council of Teachers of English.
- Messick, S. (1988) The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer (Ed.), *Test validity* (pp. 33-45). Hillsdale, NJ: Erlbaum.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Moss, P. (1994a). Can there be validity without reliability? *Educational Researcher*, 23(2), 5-12.
- Moss, P. (1994b). Validity in high stakes writing assessment: Problems and possibilities. *Assessing Writing*, 1(1), 109-128.
- Office of Institutional Research, Assessment and Planning at Kingsborough Community College. (2007). *2006-07 performance management report*. Retrieved June 2, 2006 from http://www.kbcc.cuny.edu/irap/pdf/pmp_highlights/pdf.
- Rose, M. (1985). The language of exclusion: Writing instruction at the university. *College English*, 47(4), 341-359.
- Rose, M. (2001). Narrowing the mind and page: Remedial writers and cognitive reductionism. In K. Halasek & N. P. Highberg (Eds.), *Landmark essays on basic writing* (pp. 23-55). Mahwah, NJ: Hermagoras Press.
- Scharton, M. (1996). The politics of validity. In E. M. White, W. D. Lutz, & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, and practices* (pp. 53-75). New York: Modern Language Association.
- Schmidt, B., Badillo, H., & Brady, J. V. (1999). *The City University of New York: An institution adrift*. New York: Report of the mayor's advisory task force on the City University of New York.
- Shaughnessy, M. (1977). *Errors and expectations: A guide for the teacher of basic writing*. New York: Oxford University Press.
- Shepard, L. (1993). Evaluating test validity. *Review of Research in Education*, 19, 405-450.
- Slevin, J. F. (2001). *Introducing English: Essays in the intellectual work of composition*. Pittsburgh, PA: University of Pittsburgh Press.
- Sternglass, M. S. (1997). *Time to know them: A longitudinal study of writing and learning at the college level*. Mahwah, NJ: Erlbaum.
- White, E. M. (1996). Response: Assessment as a site of contention. In E. M. White, W. D. Lutz, & S. Kamusikiri. (Eds.), *Assessment of writing: Politics, policies, and practices* (pp. 301-304). New York: Modern Language Association.
- Williamson, M. (1994). The worship of efficiency: Untangling theoretical and practical considerations in writing assessment. *Assessing Writing*, 1(2), 47-73.
- Williamson, M., & Huot, B. (2000). Literacy, equality, and competence: Ethics in writing assessment. In M. A. Pemberton (Ed.), *The ethics of writing instruction: Issues in theory and practice* (pp. 191-209). Stamford, CT: Ablex.