

5-1-2014

(Re)Envisioning Placement for 21st Century Writing Programs

Jessica Nastal-Dema

University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Higher Education and Teaching Commons](#)

Recommended Citation

Nastal-Dema, Jessica, "(Re)Envisioning Placement for 21st Century Writing Programs" (2014). *Theses and Dissertations*. 736.
<https://dc.uwm.edu/etd/736>

This Dissertation is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

(RE)ENVISIONING PLACEMENT FOR 21st CENTURY WRITING PROGRAMS

by

Jessica Nastal-Dema

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in English

at

The University of Wisconsin—Milwaukee

May 2014

ABSTRACT
(RE)ENVISIONING PLACEMENT FOR 21st CENTURY WRITING PROGRAMS
by
Jessica Nastal-Dema

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Professor Charles Schuster

Writing assessment has continued to gain prominence within educational settings and public discourse in the United States throughout the past century. Placement into First-Year Writing, however, is consistently ignored in the scholarly literature, despite its central role within writing assessment and the university. This dissertation argues that placement is central to all students' university experiences, and deserves more attention. Placement is at the heart of composition: it affects each student, each instructor, each writing program, each institution. It significantly influences retention, instruction, budget, and even national reputation, since student retention and graduation rates are key factors in national rankings such as those published by U.S. News and World Report. Placement communicates what the composition program and university value about writing to students, instructors, academic departments, and the public. It influences what high school and college English departments teach and how teachers, curricula, and programs are administered. It offers incoming students their first experiences with the university's expectations for writing, and has irrevocable consequences beyond a student's first semester. Yet placement is often invisible and ignored on campuses and in the scholarly literature, which this dissertation aims to change.

Research from educational measurement and academic communities over the past century has been positioned as an argument, not a discussion. Consequently, placement methods developed from each community have adhered to distinct criteria: efficiency and

reliability from psychometricians, pedagogy and validity from academics. This dissertation examines those criteria within four prominent methods of placing students into their required First-Year Writing courses: multiple-choice tests of grammar and usage, holistically scored essays, Directed Self-Placement, and Automated Essay Evaluation. Because writing assessment is best when it is attuned to local curricula and student bodies, I do not argue for one “gold standard” placement practice. Rather, I argue for a *dialogic perspective* of placement that considers placement in relation to the university – its students, its location, its mission, its writing values. I present suggestions for how such a perspective might be used to influence equitable practices in the twenty-first century, and to situate placement as a viable site of inquiry.

Keywords: writing assessment, validity, writing program administration, composition pedagogy, writing studies research, directed self-placement, automated essay evaluation

© Copyright by Jessica Nastal-Dema, 2014
All Rights Reserved

For Xhelal

TABLE OF CONTENTS

Preface	1
Chapter One: Why Placement Matters	4
Introduction: Assessment Despair	4
Composition and Writing Assessment	7
The Historical Significance of Writing Assessment	9
Writing Assessment, Pedagogy, and Programs	15
Why Placement Matters	18
What Placement Communicates	20
Structure of the Dissertation	24
Chapter Two: Methodology	27
The Historical Situation of Placement	28
The Role of Context	30
Classical Test Theory: Reliability & Validity	35
Validity and Placement into Composition	42
Writing Assessment in Composition Studies	47
Site-based	50
Locally-controlled	53
Context-sensitive	56
Rhetorically-Based	58
Accessibility	59
(Re)Envisioning Placement	61
Chapter Three: The Placement Trifecta: Quick, Cheap, and Accurate	63
The Labor of Placement	66
Multiple-Choice Tests of Writing Ability	69
Resistance to the Tests	72
Objectivity and Standardization	73
Validity	76
Consequences	78
Holistic Scoring of Essays for Placement	81
Rubrics	83
Reliability	86
The Prominence of Writing Assessment in Higher Education	88
Chapter Four: In Search of Better Placement	92
Strengths of Directed Self-Placement	93
Implications of Directed Self-Placement	99
Student Satisfaction	106
Placement through Automated Essay Evaluation (AEE)	108
Implications of AEE	113

Chapter Five: (Re)Envisioning Placement in the Twenty-first Century	128
An Efficiency Model of Placement	128
Methods of Placement	131
Language Difference in the University	136
Placement in FYW	142
Neo-empiricism and Dialogue	147
A Dialogic Perspective of Placement	150
Language and Power: Validity	152
Language and Power: Numbers	155
Conclusions and Implications	159
Challenges to Placement Decisions	162
Next Steps	164
Works Cited	166
Appendix A	184
Appendix B	187
Appendix C	189
Appendix D	191
Appendix E	194
Appendix F	198
Appendix G	201

LIST OF FIGURES

1: Validity Definitions from Educational Measurement Specialists	39
2: Validity Definitions from Composition Scholars	40
3: Brian Huot's Principles for a New Theory and Practice of Writing Assessment	48
4: Various Purposes of Placement	67
5: Traditional Criteria for Placement Practices	133
6: University Mission and Vision Statements	141
7: Modern Criteria for Placement Practices	145
8: Divergent Ideologies	150
9: Chris Gallagher's Bureaucratic-Institutional Model of Authority	151
10: Example of Translated Scores	158

ACKNOWLEDGMENTS

To my students at the University of Wisconsin—Milwaukee and Saint Louis University (Madrid) for inspiring and challenging me.

To Chuck Schuster for mentoring and encouraging me.

To Ariana Pantaleo, Ingrid Nordstrom, Nic Learned, Katie Morrissey, Rachael Sullivan, and Lauren Isaac for support, goodwill, and cheer.

To my parents, Bob and Patti Nastal, for my roots and my wings.

Preface

English 101, Spring 2010

It's the second semester that I'm teaching first-semester composition in a new university but my third year teaching writing. I meet with students individually to discuss their work, and many of the conferences reveal that a majority of this class has already failed English 101 at least once, and sometimes twice. The few students who haven't failed have taken English 095, the developmental writing course on campus that carries no credit toward graduation. Most students resent being in the class – they still need to complete English 102 with a “C” or better to fulfill their General Education Requirement, and it seems like it will never happen.

As the semester continues, students share their frustration with each other and with me. Somehow this makes the class easier to face for me, and the writing easier to do for them. One student, someone who is always in class and eager to participate, confesses she thinks she's a bad writer. We ask her to tell us more. She was in English 095 last semester. We ask why. She doesn't know. She always thought she was a good writer. No one explained her score on the placement test. Her teacher last semester didn't understand why she was in the class, either.

English 095, Fall 2012

It's the second semester I'm teaching English 095, Fundamentals of Composition, and the first semester I'm coordinator of the course. I've developed a curriculum based on our First-Year Writing Program's learning goals and outcomes. I understand students enter the course based on scores they earned on a placement test that places students relatively accurately into or out of the FYW sequence. In fact, I contact students who

have enrolled in the wrong course, based on their placement scores, and direct them to enroll in the correct course.

The curriculum emphasizes critical reflection of student writing. The first assignment, similar to those in English 101 and English 102, asks students to read and reflect on guiding principles of the FYW program. Students this semester confess learning disabilities, reading and writing difficulties, and resignation about being in a developmental writing course. They cautiously express hope that their teacher knows what she's doing and can help them pass the class. They don't mention "success."

Later in the semester, students reach out to me individually. They ask why their writing is "great!" in my class but receives failing grades in other classes, why other teachers take points off for glimmers of their linguistic differences. One student spends class being simultaneously interested and bored. We read "Students Right to Their Own Language," and she is outraged. She asks what the point of this policy from 1974 is when students in 2012 have to take a grammar test to determine their first required writing course in the university. She angrily continues: Why is she in this class when she has been in four years of high school Honors English? When she used to enjoy writing? When her high school peers and instructors think she is a good writer?

Prelims and Dissertation, Fall 2011 – Spring 2014

I am working on my preliminary exam reading list and proposal this fall, and have decided Writing Assessment should be one of my research areas. I pass prelims and begin working on a project that considers the assessment loop, including placement. I write my dissertation proposal as I teach English 095 in Fall 2012 and focus exclusively on placement. I use the project to understand why and how my students place into

developmental writing; to understand the history of placement; to understand the results and repercussions of placement practices and policies; to help create a better writing program here or a similar urban public school or really any university. I continue to work on the project in order to rethink what it means to use writing assessment to improve teaching and learning;¹ to reclaim assessment, for instructors and writing program administrators;² to rearticulate placement in terms of its position within higher education;³ to reframe placement to acknowledge its dynamic history and its dialogic nature within a range of fields like composition, rhetoric, educational measurement, and admissions;⁴ to re-envision what an ethical, conscionable theory of placement might look like and what a method of placement can do to improve teaching, learning, and higher education experiences. From out of these experiences and concerns emerges this dissertation, a study of placement and its role in higher education.

¹ Wilson, Maja. *Rethinking Rubrics in Writing Assessment*. Portsmouth, NH: Heinemann, 2006. Print.

² Gallagher, Chris W. *Reclaiming Assessment: A Better Alternative to the Accountability Agenda*. Portsmouth, NH: Heinemann, 2007. Print.

³ Huot, Brian. *(Re)Articulating Writing Assessment for Teaching and Learning*. Logan, UT: Utah State University Press, 2002. Print.

⁴ Adler-Kassner, Linda, and Peggy O'Neill. *Reframing Writing Assessment to Improve Teaching and Learning*. Logan, UT: Utah State University Press, 2010. Print.

Chapter One: Why Placement Matters

Introduction: Assessment Despair

Writing assessment is undoubtedly gaining prominence within public discourse about education. Recent changes to the SAT make this exceedingly clear. On March 5, 2014, the College Board announced significant revisions to the SAT, which had last been updated in 2005. Modifications include a closer alignment between the test and the high school curriculum, presumably linked to the Common Core State Standards; a smaller range of mathematical concepts being tested for deeper comprehension; texts from foundational documents in American history and science; a simplifying of the vocabulary section; and a now-optional writing test focused more on critical thinking and analysis than length and multi-syllabic words. These changes have inspired conversations among policymakers, educators, journalists, and the general public, perhaps because most of us in the United States of America can relate to the high school, large-scale, high-stakes testing experience. A recent Google search for “new SAT” resulted in nearly 1.2 billion hits, with conversations happening in education publications and websites, including *Inside Higher Ed* and *The Chronicle of Higher Education*; major news sites in the US and UK; local newspapers and websites; magazines like *Time*, *The Atlantic*, *The New Republic*; and even technology-centered websites like *PC World*. Based on this response alone, we can reach one indisputable conclusion: college admissions and placement into FYW matter.

Despite – or perhaps because of – its notoriety, writing assessment is also becoming increasingly important within writing programs, in the field of Rhetoric and Composition, and across universities nationwide. With increased governmental and

corporate participation in value-added learning, backwards-design, testing-based accountability, and outcomes-based curricula initiatives nationwide comes what Chris Gallagher has labeled “assessment despair.” He explains:

[the accountability agenda] has hijacked assessment...turning it from an instructional tool into a policy tool. Assessment has become the surveillance device of choice for politicians, policymakers, and administrators charged with maintaining “an efficient and accountable bureaucracy” (Huot). And when high stakes are attached, it becomes a weapon wielded against teachers and students who don’t measure up, who fall outside the normative, who aren’t deemed to be repaying the investment made in them. (Gallagher *Reclaiming* 56)

It is no surprise that with the increased presence of writing assessment on campuses and in public discourse comes an increased wariness about it. What Gallagher comments on, and what teacher-scholars in composition respond to, is assessment used as a weapon, as a way to decrease funding, eliminate certain courses, and increase class sizes. Writing assessments that threaten access, teaching, and learning for students have certainly contributed to today’s assessment despair.

Contrastingly, good writing assessments help writing program administrators (WPAs), instructors, and student writers themselves to improve the teaching and learning that takes place in our universities. As programs conduct assessments, for example, they examine which moments in the curriculum seem to help improve student writing, or which concepts students appear to struggle with by the end of a semester. They use this information in many ways to improve each aspect of a writing program, for instance, to

revise the curriculum, re-work the course learning outcomes, offer opportunities for professional development so instructors can better meet their students' needs. Although most faculty find it bothersome and view it as a move toward surveillance, they also understand the value of assessment.

Placement, however, is less visible and often takes place “behind the scenes” (Cushman 269). To some scholars, placement is a necessary and democratizing tool. It is necessary to determine which students could benefit from the most instructional support and intervention, and it is democratic because students are not prevented from admission based on their writing abilities. Edward M. White claims placement tests and the developmental writing programs some students place into “[serve] to help underprepared students succeed instead of washing them out” and continues, “these are the students for whom required placement and the required freshman course are necessary, for they are most in need of guidance and support...” (“Importance” 76-77). Presumably, placement testing helps those most at-risk students because it identifies and targets them for developmental writing courses, which are typically smaller than mainstream courses so students can have increased interactions with and instruction from their instructors. Many others, including Mary Trachsel and Carolyn Marvin, remain unconvinced that writing assessments in their current state can be a site of positive change. They claim high-stakes assessments are all too often mechanisms that propagate white mainstream middle-class values and that penalize students of color, linguistic diversity, and lower socioeconomic status. Writing teacher-scholars state, “educational tests are more apt to function as mechanisms that enable an educated elite to impose exclusive standards upon academic aspirants,” and even more dramatically, that large-scale standardized tests are

“instruments of social tyranny” (Trachsel 22; Marvin qtd. in Trachsel 22). Others view these tests – including placement into composition – as instruments of exclusion:

In the current mean-spirited political climate, I doubt whether we serve “new students” [typically students of diverse cultural, ethnic, linguistic, and economic class backgrounds] well by using mass examinations to segregate them into classrooms that can readily be identified as remedial or special. (Crowley “Response” 90)

It is within this mean-spirited political climate and environment of assessment despair that I write to confront Crowley’s question– *Can we serve diverse student bodies well through placement practices?* – and to answer, *I think we can.*

Composition and Writing Assessment

Composition scholarship is perhaps best known for its merging of theory and practice. Teacher-scholars are dedicated to finding ways to uncover helpful methods that encourage the development of student writing while remaining cognizant of current theoretical and scholarly research. As a subfield of composition studies, writing assessment must also be mindful that its theories and practices treat students fairly and ethically, no easy matter given student diversity and programmatic differences in high schools and colleges. From the perspective of writing program administrators in higher education, for example, assessment is a key component in the mission to create coherence within writing programs from curricula to evaluation to the professional development of composition instructors. The tradition begins with Paul Diederich’s work at Educational Testing Service (ETS) in the 1970s, which encouraged holistic scoring of writing samples in addition to multiple-choice tests to fairly evaluate student work; it then moved to

Richard Lloyd Jones, Lee Odell, and Charles Cooper's development of primary trait scoring in the 1980s, and portfolio assessment in the succeeding decade under Peter Elbow and Pat Belanoff's influence.

Each of these efforts at assessment emerged from specific understandings of theory, practice, and a commitment to have students assessed accurately and fairly. Mya Poe claims, "[t]he early history of assessment reminds us that design, interpretation, and consequence are always intertwined" (272), which has been especially evident in writing teacher-scholars' work in assessment. As writing instructors encouraged the development of student agency in their writing, for example, leaders like Edward M. White began to draw the composition field's attention to ways in which holistic scoring actually discourages thoughtful, independent writing (see his scathing "My Five-Paragraph-Theme Theme"). Michael Williamson, Brian Huot, Peggy O'Neill, Susanmarie Harrington, and Bob Broad became pillars of the next "wave" of writing assessment, to use Kathleen Blake Yancey's term, with increased attention to programmatic assessment. This movement has a concentrated focus on the rhetorical aspects of assessment. That is, these teacher-scholars highlight how the best – and most ethical – writing programs emphasize that both writing and assessment are locally situated and driven. Yancey explains, "writing assessment is now [1999] constructed as a rhetorical act: as its own agent with responsibilities to all its participants" ("Historicizing" 501). The third wave embraced a new era in writing assessment that emphasizes "how well assessments support best practices in teaching and learning" (Broad *Value* 137). In the fourth wave, our current era, practitioners collaborate across disciplinary and geographic divisions to

respond to an exigence beyond local concerns while working to address rhetorical concerns in assessment and in student writing (Yancey “Rhetorical” 477).

It is from this current wave of writing assessment that my dissertation emerges. I use terms central to both psychometrics and writing studies – validity and context – and discuss how writing program administrators might re-envision them for twenty-first century writers and programs, with a particular focus on one of the more neglected aspects of assessment: placement into first-year writing. I pay careful attention to the scenes of writing – curricular and physical – to determine just how we can use assessment formatively, to strengthen our programs and our students’ experiences with writing in higher education. Ultimately, I approach placement from a dialogic perspective that builds on histories of writing assessment and composition to (re)envision placement.

The Historical Significance of Writing Assessment

Writing pedagogy emphasizes the centrality of assessment to the act and instruction of writing; we know good writers are also good self-assessors, and encourage students to take part in various peer- and self-review to instill habits that can translate into their own writing. Instructor feedback is built on a formative view of assessment, on the belief that our intervention can help students re-view their writing and revise it to better address their purposes and audiences. The use and history of writing assessment outside of the classroom, however, is more complicated. From recent discussions in public discourse and popular media, the weight of writing assessment is significant. It goes beyond affecting individual instructors’ assignments and student submissions. It influences local and regional curricula, educational policies, and even school funding. Writing assessment has also historically been used in the United States to make decisions

about test-takers' intelligence and has subsequently influenced federal laws. In a recent special issue of *Research in the Teaching of English* dedicated to diversity and international writing assessment, guest editor Mya Poe briefly summarizes its history:

Assessment, including writing assessment, has never been about homogenous populations of students nor about isolated practices. Because what mattered to early test developers was the development of instruments to measure innate ability, not the impact of assessments on test populations, variations in populations and contexts were ignored. But the consequences of test use – following Galton's theories of eugenics – were disastrous (Zuberi, 2001). In the United States, differences in test scores were used as evidence of racial inferiority. (272)

Poe's focus is on the consequences of writing assessment, with particular attention to historically marginalized groups of people and students. Her summary points to ways in which writing assessment has been used on national levels with terrible, racist consequences. Connections between testing, literacy, and race were made possible by the advent of intelligence testing by Alfred Binet at the turn of the twentieth century led to new perceptions of cognitive ability. Binet's tests were adapted to United States Army recruitment efforts for World War I, where literacy became conflated with intelligence (Patelis; Karabel; Lemann; Elliot). That is, when recruits scored poorly on the exams, they were perceived to be unintelligent and even intellectually disabled – not necessarily illiterate, and with little consideration for the lack of educational opportunities for working poor and minority recruits. Furthermore, a perceived lack of intelligence has also historically been aligned with a flawed character, which has led to harmful laws and

social, including educational, practices (see Hull et al. and much of Mike Rose's work). Despite the problems associated with intelligence testing, the methods were adapted for use in higher education.

Carl Brigham, who worked on developing the Army Alpha and Army Beta tests, was instrumental in adapting them into the Scholastic Aptitude Test (now known simply as the SAT). Entrance requirements varied widely by institution, which the nationally-used exam helped to standardize. High schools were consequently better able to create curricula that prepared students for higher education, which was increasingly important in an era that emphasized the production of a managerial class rather than a ruling elite (Patelis). A standardized approach to university admissions was also becoming increasingly appealing due to the explosion of students. The number of high schools increased 200 times between 1869 and 1880 (Patelis); university bodies were more diverse than ever before. In *On a Scale: A Social History of Writing Assessment in America*, Norbert Elliot provides a chronological examination of shifts within writing assessment. Elliot's analysis offers an excellent introduction to the field as it clearly articulates how writing assessment has been inextricably tied to various social – and political – developments within the United States. In the post-World War II era with an exponentially larger university student body, entrance exams needed to be efficient in terms of cost and speed, and multiple-choice tests fit those requirements. The tests, including the SAT, could be widely distributed and quickly scored to provide information about the test taker.

The increase in students applying for admission, in the Midwest in addition to the East Coast institutions, alongside the development of the SAT shifted admissions

priorities. Students in the “Big Three” universities – Harvard, Yale, and Princeton – historically demonstrated their abilities in a range of subjects through writing (Broome; Elliot; Karabel). As the number of students increased, however, it became difficult to read scores of essays as universities had done in the nineteenth century. Furthermore, as psychometricians dove into intelligence and aptitude testing, it also became evident than the reliability of essay tests, especially the agreement between different readers, was weak:

The reliability problem is perhaps best illustrated by a simple example. In 1961 a study was conducted at the Educational Testing Service in which 300 essays written by college freshmen were rated by 53 readers representing several professional fields (French 1962). Each rater used a nine-point scale. The results showed that none of the 300 essays received less than five of the nine possible ratings, 23 percent of the essays received seven different ratings, 37 percent received eight different ratings, and 34 percent received all possible ratings. It was clear from this study that the score received was to a large degree dependent upon which expert happened to be doing the scoring. (Breland 1)

In the 1961 study, a third of student essays received nine different scores. Hunter Breland’s findings are echoed by Paul Diederich, who found 94% of essays received at least seven different scores (Diederich). Upon closer examination, the variance in scoring becomes evident:

The literature indicates that essay tests are often considered more valid than multiple-choice tests as measures of writing ability. Certainly they

are favored by English teachers. But although essay tests may sample a wider range of composition skills, the variance in essay test scores can reflect such irrelevant factors as speed and fluency under time pressure or even penmanship. Also, essay test scores are typically far less reliable than multiple-choice test scores. (Cooper 1)

Despite standardized scoring rubrics, human readers have personal preferences and beliefs about what constitutes good writing (and as a long line of my teachers can attest to, if a writer's handwriting is illegible, it is impossible to determine their writing abilities). Because of the difficulty in reaching consensus among readers, the subsequent task in writing assessment became to establish which methods could best – most accurately – reflect students' writing abilities. If readers could not agree on how to evaluate student writing, if essay tests proved to be unreliable, perhaps multiple-choice tests could provide better information.

Multiple-choice tests were established to be reliable and to provide helpful information regarding students' verbal and mathematical abilities. Subsequently, the SAT was firmly in place to standardize admissions practices nationwide (Lawrence et al.). But conversations about writing assessment in that era are echoed today: What is writing ability? How can it be measured? How can a test of writing be measured reliably? How can it be valid? How can tests of writing ability be fair to students? How can a writing test demonstrate the breadth of students' preparation for higher education? As the SAT gained ground, testing experts began to question the merits of an entrance exam that heavily emphasized abilities in English alone or at the expense of other subjects – mathematics, sciences, history, etc. – to garner admission to the university, just as writing

instructors resisted exams of writing that did not require students to write. In fact, writing teachers were not convinced that writingless tests *could* portray their students' writing abilities, and viewed the tests as inherently flawed:

Historically...the objective test has been severely criticized on the grounds that it presents the examinee with a task which is artificially oversimplified. It has been charged that the examinee is inadequately measured when he is required merely to choose his answer from among a number of answer-choices which are set down for him. (Huddleston 165)

The tests were also – and continue to be – criticized because they “emphasized mechanical skills at the expense of style and quality and because ‘pupils may respond correctly in objective tests to items which they do not use correctly in their own expression’” (Huddleston 166).

The 1960s and 1970s saw a developing field of Composition, the Civil Rights Act, barriers to education breaking for African-American and Hispanic students, and renewed questions about testing students' writing abilities. The SAT shifted its focus to consider its consequences on diverse student bodies who could no longer be ignored (Lawrence et al.). Writing teacher-scholars developed new methods of teaching writing and of evaluating that writing. Those shifts in approaches to teaching and assessing writing at the college level, in the context of a classroom or writing program, have been well documented in composition scholarship. The changing makeup of student bodies enrolled in universities and colleges appears to have given rise to renewed concerns about their academic preparation – a recurring theme any time admissions policies become more open. From my research, it seems that admissions tests suggested an increasing

number of students were severely underprepared for writing at the college level, which echoes Brereton's account of the development of "English A" at Harvard, and which helped to usher in the Basic Writing movement (see Shaughnessy). The development of placement testing itself, as an entity separate from admissions testing, is less apparent in the scholarly literature. It is clear that writing teacher-scholars became more concerned not only about how students were assessed at the end of a course, but also about the processes by which students entered the course. This is evident in reviewing the changing approaches to entrance into FYW, from multiple-choice tests to short essays to Directed Self-Placement to Automated Essay Evaluation (all of which this dissertation analyzes). It is less apparent what happened to placement into writing courses during the same era, how it separated from admissions testing as well as FYW programs, however, which I plan to examine in a future study based on archival research.

Writing Assessment, Pedagogy, and Programs

University entrance exams are just one area where writing assessment has been prominent for the past century; in fact, Yancey claims, "writing assessment has always been at the center of work in writing" ("Historicizing" 483). In her view, writing assessment is the central aspect of the administering, teaching, and even learning about writing. It is ubiquitous and occurs in places across the university including admissions tests, proficiency and exit exams, and writing classes. But as writing assessment has "permeated the entire institution of composition," it has also been invisible ("Historicizing" 483). That is, it has often taken place outside the composition classroom or writing program, leaving the fate of students in the hands of anonymous testing experts. Although it would be difficult to claim assessment is "invisible" in 2014, one

area of writing assessment still is: placement. Placement into composition courses occupies the position of assessment Yancey claims existed in the first wave, 1950-70. It is ubiquitous – each student must experience it; each instructor must work with the consequences of it – yet invisible as it is often removed from the control of writing programs.

A major goal of the third wave of writing assessment, rolling into the fourth wave, is to reclaim writing assessment from the accountability agenda (for Gallagher) and from the dominant hands of psychometricians. Brian Huot, in his foundational text for both writing assessment and composition studies, *(Re)Articulating Writing Assessment*, maintains that it is essential to align writing assessment and curricula. Otherwise, we will create writing programs and courses with no clear direction or sense of continuity and assessments with no basis in practice. In the third wave of assessment, it is essential for writing instructors and administrators to be aware of and participate in the articulation, development, and maintenance of writing assessments. While doing so certainly allows us to create better-informed curricula, which in turn helps to improve the teaching and learning in our classes, it is also an act of self-preservation. As White reminds us, “Do assessment yourself, or it shall be done unto you,” and more emphatically:

the effects of the assessments are profound for students, teachers, and the curriculum; writing teachers must be present, informed, and vocal when decisions are made, or else crucial matters slip away from the faculty and into the hands of administrators and clerks who may know nothing about the teaching of writing. (*Assigning* 100)

By not participating in conversations about assessment, curricula are removed from the influence of those who are best prepared to lead instruction and assessment efforts.

Furthermore, people generally assess what they value, and as Huot explains, “Because assessment is a direct representation of what we value and how we assign that value, it says much about our identities as teachers, researchers, and theorists” (Huot, *(Re)Articulating* 11). The things we assess are public declarations of what we care about. If compositionists do not participate in writing assessments, and they are instead in the hands of an interdisciplinary university committee (for instance), we have no guarantee students will be required to complete meaningful writing assessments. There is no guarantee students will even be assessed on what our discipline values, as articulated by the WPA Outcomes Statement: writing that demonstrates rhetorical awareness and knowledge; critical thinking, reading, and writing; evidence of various writing processes (including drafting and collaboration); in addition to the knowledge of conventions (see Appendix A). Additionally, participants in our various constituencies – students, instructors, university members, parents, potential employers – will come to believe that the writing program values arhetorical writing!

Other scholars urge writing assessment practitioners and writing program administrators to take their central role in pedagogy and assessment to the public sphere, to become immersed in conversations about their practice and scholarship “wherever they might occur” (Adler-Kassner and O’Neill 9). Linda Adler-Kassner and Peggy O’Neill argue that the only way WPAs can initiate sustainable change within writing programs is to understand public discourse about writing, work from within it, and reframe it to better

suit the needs of students, instructors, and WPAs. For these leaders in assessment and writing studies scholarship, conversations about writing assessment are

the most important discussions happening on our campuses (and even beyond them) today. They affect everything about our courses and programs – who is admitted to them; how they are taught; how students, courses, and instructor(s) are evaluated; what counts as valuable in them – in essence, everything that motivates writing instructors to do the work we do. (Adler-Kassner and O'Neill 4; emphasis original)

Because writing assessment affects every aspect of a composition instructor's and WPA's work, it is essential that we participate in conversations within our programs, departments, colleges, universities, and communities. One way we can better frame those conversations is to ensure our practices communicate our values.

Why Placement Matters

Despite its significant impact on multiple levels: for the university, department, program, course, and individuals, placement into composition is a relatively under-theorized aspect of the assessment process.⁵ At first glance, it seems like placement into writing courses only affects students who are assigned to courses that don't fulfill their needs; or perhaps instructors or WPAs who need to move students between sections. But placement is at the heart of composition: it affects each student, each instructor, each writing program, each institution. It significantly influences retention, instruction, budget, and even national reputation, since student retention and graduation rates are key factors in national rankings such as those published by U.S. News and World Report. Placement

⁵ For instance, my research resulted in no books solely dedicated to placement, and one quarter as many hits on placement compared to assessment on CompPile.

communicates to students, instructors, academic departments, and the public what the composition program and university value about writing. It influences what high school and college English departments teach in addition to how teachers, curricula, and programs are administered. It offers incoming students their first experiences with the university's expectations for writing, and has irrevocable consequences beyond the first semester students are enrolled. Yet placement is often invisible and ignored on campuses and in the scholarly literature.

In its simplest form, writing assessment evaluates student writing in terms of how well it has achieved a number of goals. We use it in formative ways, to gauge student development, and summative ways, to determine programmatic success. Placement, however, is driven by prediction. Placement is one way we determine students' proficiency, but we use it after they have been admitted to the university. It is a sorting tool, not a requisite for admission. It determines which students need the most assistance in their writing and which students are already writing at college levels. We use it to determine how well a student might do in a given course. We want students to build on their previous skills and to be challenged just enough to fulfill our learning outcomes. It is also often built on a deficit model, reminiscent of the nineteenth-century Harvard entrance exams. At the time, students were perceived to be unprepared for college work and needed to write essays to prove their readiness; this model is still in place today when we assume students cannot make accurate decisions about their writing readiness and instead must be tested to determine how they "best fit" within our often-required composition classes (White "Foreword").

Although placement is still a widespread practice, Richard Haswell and William Smith point out that it is mostly a guessing game. Students often perform better the second time they take a test, so the reliability of their scores on placement exams come into question. Placement tests that rely on multiple-choice questions about grammar and usage offer little information about the writing students actually do, so their validity is suspect. Processes that require students to make an informed decision are criticized for pushing the responsibility onto students, or for putting too much of a financial strain on writing programs and universities.

Even as writing programs and universities recognize the importance of placement, its cost and time demands have made many institutions reluctant to change what is currently in place. In an era of slashed budgets and overworked staff, the predominant feeling about placement has been summed up quite clearly by William Smith: “It seemed to work, so there was no impetus to examine it, let alone change it. The incoming students were placed into our courses efficiently and with what appeared to be tolerable numbers of errors” (“Importance” 314). Correspondingly, innovative writing programs and graduate programs in Rhetoric and Composition incorporate cutting-edge research in their curricula and end-of-semester assessment methods, but employ outdated indirect testing methods to determine student placement into composition classes, a practice that continues to influence public conversations about college students’ inability to write.

What Placement Communicates

Placement into composition does not just determine one course for a student: “In placement testing we actually decide for a student where she will be placed for the next fifteen weeks, or, perhaps even more importantly, where she will begin her college and

university writing instruction” (Huot *(Re)Articulating* 6). In each placement decision, it is essential to remember we are not only making a decision about the merits of an exam or essay; we are making decisions about a student’s future experiences in higher education. We should, therefore, not be content to simply maintain the status quo.

Drawing from the scholarship of Bob Broad and others, this dissertation is concerned with “what we really value”: how various placement practices articulate values to participants in our various communities, which include students, instructors, the university community, governmental bodies, and the concerned public. I believe it is necessary to embark on such a research project for several reasons. First, placement is the initial experience students have with a first-year writing program; it should be a positive one for students that fosters their understanding of and participation in the writing curriculum. Second, as flagship universities, it behooves us to employ innovative, ethical methods not only in our curricula and end-of-semester assessments but also in the placement processes that determine our students’ academic destiny. Third, urban-serving universities like UW–Milwaukee, a paradigmatic institution in many ways, are guided by mission statements dedicated to their diverse, urban environments. Such universities should therefore feel some pressure to move beyond standardized multiple-choice placement testing, a modality that has a reputation for being culturally specific and discriminatory toward minority students.

As most composition administrators would acknowledge, placement in composition courses is typically accomplished by requiring students to complete local, regional, or national standardized tests that at best offer indirect measures of writing ability. Such tests de-emphasize the very ability they are intended to measure. They

promise efficiency and reliability at a low financial cost, and so they have proliferated. But they have other costs: they send a strong message to high schools about college English preparation; they communicate a negative view to students about what writing is at the college level; and they misplace students whose writing is much stronger than their test-taking ability – or whose test-taking abilities are stronger than their writing. A basic premise of this dissertation is that those of us who teach and administer First Year Writing have an obligation to communicate our values at each step of the FYW experience, including placement. One central value is that placement into writing courses must involve writing. Although such a statement may appear oxymoronic, it is observed more in its lapse than its observance. I argue such indirect methods send a negative message to students and the university community.

In *WPA Outcomes Statement: A Decade Later*, the editors explain the centrality of an institution's learning goals and outcomes: "Yet, college writing programs, despite their mission statements, curriculum guidelines, and examinations of all sorts, have rarely attended to student outcomes, and it is those outcomes that provide clear opportunities for assessment and clear data for outsiders looking for accountability" (Behm et al. x). Why not align our placement processes with the learning goals and outcomes we work hard to design? Why not use placement as a way to improve the first year writing course and perceptions about its purpose to all our various stakeholders: teachers, students, administrators, parents, governmental bodies...? Doing so, to Brian Huot, "can change the public evidence about what is valuable" ((Re)*Articulating* 9). It can create coherence within our programs. And ultimately, it can position our values about student writing in a

powerful way at the local, regional, and national level by aligning with articulated learning goals and outcomes.

Throughout composition studies, we demonstrate a dedication to creating spaces that allow students to assert their agency to improve their experiences in FYW. As Asao Inoue et al. explain, “students can and should have more agency in placement processes, and when they do, they perform better in classes and are happier with those classes.” Placement is the initial step in how students “come to understand themselves and others, as writers, as constructors of knowledge and members of educational and broader communities in which they participate” (Moss “Testing” 119). If writing instructors and administrators participate more fully in the placement procedures on their campuses, we can try to represent students and their needs better. Similarly, we can guarantee better *validity* of those procedures as we design them with a clear understanding of both their intended purpose and their likely interpretations.

I believe we need to inform *all* of our stakeholders about the central importance of writing. Of treating students ethically and fairly. Of including writing for humans – not machines – in our assessment methods.

Of including writing in our placement measures.

Of creating writing programs that are *coherent*, where placement procedures, curricula, student- and program- assessments reflect – and assert – our values.

And I believe one way we can begin shifting how the public views writing is by creating placement procedures that use rhetorically-situated learning outcomes, just as our curricula and assessments do.

Structure of the Dissertation

My approach to placement, and therefore to this dissertation, relies heavily on my approach to writing pedagogy. I work from a historical approach to composition and to assessment, which enabled me to recognize the complaint that “Johnny Can’t Write” has been a recurring theme in public discourse since the famous Harvard entrance exams. This approach allows me to understand how literacy in the late nineteenth century marked intelligence, just as it continues to do in the early twenty-first century, particularly as our higher education institutions recruit increasingly diverse student bodies. Sensitivity to language is therefore central to my pedagogy as a writing teacher, drawing from the work of scholars like James Berlin, Keith Gilyard, and Mina Shaughnessy on how language, writing, and the instruction of both are inherently political acts. Ultimately, my goal is for students to emerge from my courses as confident in their abilities to shape meaning through writing, to interact with their world, and to contribute their ideas to ongoing public and academic conversations.

One goal of this dissertation is to contribute to the existing scholarship on writing assessment more generally, and to help build scholarship on placement so it too can become a well-theorized practice within writing studies. It is my contention that a historical view of assessment can help instructors and WPAs construct well-informed, adequate, and accurate methods of assessment. In this dissertation, I am particularly concerned about how to include placement in the writing assessment practices programs have come to take control of and pride in.

This dissertation reflects my dedication to writing pedagogy, to recognizing students as individuals, and to making a positive impact in the fields of composition and

writing assessment, all through the often-overlooked practice of placement. Chapter Two details my approach to analyzing current placement models, and argues for a dialogic approach. Chapters Three and Four examine indirect and direct writing assessments used to place students into composition, and analyzes them based on the foundational concepts of validity and context. Chapter Five offers a (re)envisioning of placement and argues that while there may be no one best method for all writing programs or universities to use, the dialogic perspective offered throughout this dissertation is one helpful step in creating better placement methods, which can lead to students having better experiences in higher education.

Mya Poe writes,

Writing assessment today [2014] is neither a narrow practice designed to sort writing samples neatly into categories nor the application of methods, even “best practice” methods, without regard to their effects. Today, writing assessment is best understood as (1) designing a series of strategies to increase our knowledge of a complex construct – that is, writing; (2) making meaningful decisions based on our measurement of that construct; and (3) understanding the effects of our practices on students and ourselves. In short, writing assessment is about understanding the relationships among design, decisions, and effects. (271)

As Poe urges, writing assessment in the twenty-first century is no longer simply concerned with best practices or methods; rather our focus now must be on the relationships affected and caused by writing assessment. Writing programs and universities nationwide pay careful attention to the consequences of writing assessment,

particularly evident in their mid-semester, end of semester, and programmatic assessments. The consequences of placement are often examined in light of developmental writing; however, the scholarly literature omits discussions about the complex relationships affected by placement as well as a deep understanding of how placement builds knowledge about the construct of writing. This dissertation is my attempt to begin that important work.

Chapter Two: Methodology

Placement in the twenty-first century can become a powerful way for writing programs to improve students' higher education experiences and instructors' teaching experiences – which is, or should be, the ultimate goal of *any* assessment. Mya Poe explains,

The study of writing assessment practice [...] exposes what is valued in student writing, how students are affected by assessment, and ultimately, how assessment might have a positive impact on the teaching of writing. Assessment should be transformative, and it should transform us – as researchers, teachers, and administrators – as much as it transforms our students' learning and writing. (271)

In the best circumstances, writing assessment can be a generative experience for its participants, which I explore further throughout this dissertation. Writing pedagogy depends on a formative view of assessment, where the evaluation of and reflection on student writing by instructors, peers, and self can lead to increased awareness of writing processes, enhanced self-efficacy, and improved writing. Best practices within writing assessment mirror that dedication to improvement, which is demonstrated in programmatic assessments, for instance, that evaluate how well a curricular change prepares students to fulfill learning outcomes, and in professional development opportunities where instructors and administrators meet to discuss their practices and experiences. Because students' encounters with placement are so central to their initiation to the university, we ought to take that opportunity to communicate our most central

values. As Susanmarie Harrington argues, placement is more than a decision about coursework for students. It

is most students' first contact with the theory and practice of first-year writing programs, and we would do well to make that first contact as inviting and theoretically sound as possible. To do so, we need to think less about placement as mechanism and more about placement as an opportunity to communicate. Placement is perhaps the first part of our programs that communicates to students. (12)

To fulfill its promise, placement must draw on a dialogic perspective, one that recognizes the interplay of values and perspectives at work. A dialogic perspective of placement acknowledges, interacts with, and builds on the different entities that affect and are affected by placement. First, a dialogic perspective calls on teacher-scholar-practitioners to recognize how placement demonstrates the interrelatedness of various university spaces. Second, a dialogic perspective requires writing assessment specialists to acknowledge the very heteroglossic nature of placement. This chapter articulates a dialogic perspective of placement by analyzing how it functions locally, within university mission statements, admissions policies, and writing programs; and nationally, through the familiar terms of *validity* and *context* within the fields of composition studies and educational measurement.

The Historical Situation of Placement

While the role of placement into composition has shifted throughout the evolution of higher education in the United States, the practice itself has been a constant. Edwin Cornelius Broome in *A Historical and Critical Discussion of College Admission*

Requirements, for instance, traces how admissions testing reflected the concerns of the ruling classes. In the seventeenth century, the first colleges focused on Latin, Greek, and preparing male students for the clergy, for instance. By the nineteenth century, admissions requirements expanded to include subjects like Arithmetic, Grammar, Rhetoric, and Composition – in part to determine how well students’ secondary education had prepared them for the university. By Jerome Karabel’s account in *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*, these admissions methods were useful to maintain high standards, particularly when it was more fashionable to be a better partier than student in the late nineteenth century! Ultimately, as Norbert Elliot makes clear in *On a Scale: A Social History of Writing Assessment in America*, the assessment of writing was an essential aspect of college admissions for decades. Since the mid-twentieth century, large-scale, multiple-choice, standardized admissions exams have become the norm in writing assessment, in part to meet the demand of the nation’s expanding student bodies. While students are required to provide a writing sample in these admissions exams, they are also often required to demonstrate their writing proficiency after they have been admitted to determine with which writing course they should begin their academic career.

It is clear from this brief overview that the evaluation of student writing has not always been in the purview of English departments or writing programs; rather, it has long been used as a method to determine students’ preparation for higher education in general. To fully understand the role of placement in the twenty-first century, therefore, we must certainly consider its role in sorting students into composition courses, but we must also consider the relationship between placement and admissions practices, between

placement and the university. The guiding belief throughout this project is that a careful analysis of current practices can help writing program administrators make better-informed decisions about their placement methods, and can help make better placement methods in general. By drawing on values central to writing studies, educational measurement, and open-access admissions policies, I believe we can strengthen our twenty-first century writing programs – and, in turn, improve the teaching and learning of writing.

The Role of Context

Within the field of Rhetoric and Composition, *context* holds a central role. In terms of writing, we understand the importance of discourse communities: “the now commonplace understanding that writing is done by and within groups of individuals who are linked together by shared discursive practices” (Vandenberg et al. 172-73). The first concept listed in the WPA Outcomes Statement is “Rhetorical Knowledge,” which requires students to “Respond appropriately to different kinds of rhetorical situations” and “Use conventions of format and structure appropriate to the rhetorical situation.” Rhetorical awareness is key to what many writing program administrators and instructors expect their students to learn after one year of composition instruction. Sidney Dobrin, ushering in the writing as ecology movement, states, “writing cannot be separated from place” (13), while Thomas Kent in *Post-Process Theory* claims, “writers are never nowhere” (3). Some of the most lasting books within composition studies are about a writer-scholar-teacher’s context: *Voices of the Self*, by Keith Gilyard; *Bootstraps*, by Victor Villanueva; and *Lives on the Boundary*, by Mike Rose. As a field, composition studies values these narratives situated firmly in specific settings (the New York City of

an African American boy, the Washington of an Americanized son of Hispanic immigrants, and the LA of a working class Italian immigrant family) – as evidence, each writer has won the Conference on College Composition and Communication’s Exemplar Award.⁶

In *Representing the “Other,”* Bruce Horner and Min-Zhan Lu advocate for recognizing the role of context, particularly as discourse communities and their “interactive cultural forces” both influence and are influenced by individuals (38). More specifically, “the particular social conditions of a given time and place can exert pressures on an individual to act not only in ways that maintain and renew existing ways of distributing and organizing social power but also to transform those ways” (xiii). For Horner and Lu, these pressures exist in structures like public schools, where local and federal governmental bodies influence curricula and working conditions. In those same spaces, however, diverse cultural forces are able to interact and have the potential to transform each other. As Mike Rose and Victor Villanueva demonstrate in their education narratives, students are often positioned by others as seeking to become part of the mainstream, typically represented by white, middle-class values. The changing makeup of composition classes – African American and Puerto Rican students in 1970s New York; Latino, Hmong, and Chinese students in 2014 Milwaukee – challenges those homogenized values and cultural forces. Our classroom discourse communities become multi-vocal and multilingual, and can help compositionists develop new standards that

⁶ “The CCCC Executive Committee presents, as occasion demands, the CCCC Exemplar Award to a person whose years of service as an exemplar for our organization represents the highest ideals of scholarship, teaching, and service to the entire profession. The Exemplar Award seeks to recognize individuals whose record is national and international in scope, and who set the best examples for the CCCC membership.” See <http://www.ncte.org/cccc/awards/exemplar>

reflect these voices. Placement practices ought to recognize the cultural, racial, and linguistic diversity present in our classrooms.

In the field of writing assessment, the location of writing, assessment, and students has led a number of writing programs nationwide to develop their own individual assessments and placement procedures (see the discussion of Brian Huot's guidelines later in this chapter). Bob Broad's work here has been instrumental, stemming from his monograph, *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Broad writes in response to traditional rubrics that rate student writing based on concepts like the 6+1 Trait Writing framework, which includes ideas, organization, voice, word choice, sentence fluency, conventions, and presentation (x). He explains,

Such concise lists of criteria may have adequately served the needs of writing assessment for forty years by making judgments simple, quick, and agreeable. As a guide to how texts are actually composed, read, interpreted, and valued, however, I propose that traditional rubrics are [...] dangerously unsatisfactory for purposes of contemporary rhetoric and composition [...] (Broad *Value* x)

Broad's work is in line with the basics of grounded theory in its attention to local sites of writing and assessment:

Relying on pre-established rubrics might force the assessment of areas not relevant to a department. Such rubrics allow for comparison across multiple groups, but do not express key components of writing that are major specific, or even department specific. (Migliaccio and Melzer 85)

Rather than limit the evaluation of student writing to a short list of vague criteria, Broad argues for *dynamic criteria mapping*, where instructors work together to discuss actual samples of student writing and to determine what they value about the writing. This “organic” or “locally grown” writing assessment is developed in direct response to the students, instructors, curricula, and relationships within a specific site, and Broad’s work documents how DCM has been used in a range of institutions (see *Organic Writing Assessment*). In his social history of writing assessment, Norbert Elliot concludes the book (another recipient of the CCCC Outstanding Book Award) by explaining, “By now, it should be clear that I believe that the history of writing assessment is best understood, in Burke’s terms, as an agent-agency ratio occurring within a specified scene,” and by citing examples of how context influenced shifts throughout writing assessment (337). Elliot also provides examples of how wars have shaped writing assessment, beginning with early aptitude tests that tied intelligence to literacy – and race – for the U.S. military. For instance, composition studies is well aware of how the civil rights movement influenced CUNY’s open admissions policies, which led to position statements like the NCTE “Students’ Right to Their Own Language” in 1974. Elliot’s take on the role of “specified scenes” considers how such movements affected writing assessment procedures and policies, including on federal and large-scale levels.

Embedded in even the earliest discussions about writing assessment is the claim, “it must be obvious that a test of composition skill cannot equally be valid in all colleges as a predictor of performance in freshman English” (Godshalk et al. 3). Ed White echoes the same finding some thirty years later: “Each college and university should consider the question of placement in light of its own student body and its own writing curriculum”

(*Assigning* 46). White's statement, in my view, is incontrovertible: placement must be considered within the context of a university's own students and own curricular goals. But I want to expand this view of context: for me, context is also about the actual site of placement exams, the cultural, physical, even (dare I say) the emotional and psychic site of placement. In practice and in the scholarship, context is often limited to how students understand their role as a writer in a certain setting; for instance, as a writer in Milwaukee or enrolled in a First-Year Writing course with a curriculum built on rhetorical analysis. Scholarship on writing studies and pedagogy, however, underscores the significance of "relations, locations, and positions" (Vandenberg et al.); that is, how a writer's position affects everything she does, everyone she interacts with. It is my contention that a theory of placement must go beyond the understanding of context as something that only affects a writer or a theory. Placement must consider the relationships and consequences involved within the practice, as Poe argues, particularly in relation to the specific students who enroll in our institutions and are subject to our writing assessments.

Elliot claims, "in writing assessment, context is critical" (339). I believe it is especially so for writing programs, instructors, and assessments in public urban serving universities. Over 80% of the US population lives in cities ("Factbook"). Throughout the country, urban primary and secondary schools contend with higher student attrition than the national average and lower national and state test scores. These schools are often located in impoverished areas and receive inadequate state and local funding. As a result of the poor preparation students often receive in urban schools given high enrollments, family pressures, economic stresses, and meager opportunities for self-improvement, they often enter the university with some academic deficits – but also significant language

resources that may well go unrecognized and unrewarded. To consider context, we must first acknowledge the admissions policies and missions of public, urban serving universities like the University of Wisconsin—Milwaukee. These institutions employ holistic admissions methods as a way to ensure students access to higher education, but such methods traditionally disenfranchise students outside the educational mainstream, for example, by placing a majority of students of color into remedial writing courses. I am arguing for a comprehensive view of context, one that examines admissions policies, mission statements, student populations, and results of placement practices. If we consider placement in relation to the university – its students, its location, its mission, its writing values – we can create practices that treat all students equitably and are ultimately, valid.

Classical Test Theory: Reliability & Validity

In *On a Scale*, Norbert Elliot presents a chronological narrative of how our nation's higher education institutions, testing companies, and federal government have used writing assessment since the nineteenth century. Across the various assessments are two matters of central importance to the field of writing assessment and to this dissertation: the historic tension between reliability and validity, and the role of context. As testing gained prominence in academic and public settings, psychometricians were called on to create methods that sorted test-takers accurately, uniformly, and inexpensively:

The first wave of writing assessment [1950-1970] is dominated by a single question: not the question we might expect – “What is the best or most valid measure of writing?” – but a question tied to testing theory, to

institutional need, to cost, and ultimately to efficiency (Williamson) –

“Which measure can do the best and fairest job of prediction with the least amount of work and the lowest cost?”

The answer: the reliable test. (Yancey “Historicizing” 489)

Reliability has historically had a bad reputation among writing instructors and specialists; Michael Williamson, for instance, has claimed, “The greater the reliability of an assessment procedure, the less interesting a description it provides of writing” (“Efficiency” 163). For psychometricians, reliability is essential to create a good test and to represent a test-taker’s *true score*, “the score that an individual would receive on a measure if he or she took the test a theoretically infinite number of times” (Williamson “Holistic” 16). A test is reliable when it produces consistent results. On a reliable multiple-choice test, a student will score similarly, given similar circumstances. In reliable essay tests, readers would give an essay the same score each time they read it, and different readers would score essays in the same way. Reliability certainly helps to ensure uniformity, which was perceived as essential for universities to accommodate their increasing numbers of students and increasingly diverse populations. Classical test theory depends on reliability to ensure generalizability and consistency in multiple-choice testing situations.

Some scholars, including Doug Shale and Peggy O’Neill, explain that the concept of reliability accompanied the shift from multiple-choice to writing-based assessments, without necessarily being reshaped or particularly effective as a metric in its new context. Although reliability has often been limited to *interrater* reliability, O’Neill suggests it might be able to be reframed to be better associated with validity and with increasingly

sophisticated discussions of writing assessment. In relation to writing, reliability corresponds to standardization and objectivity; it reduces or rejects the inherently subjective nature of reading and writing. It also reflects an era dominated by positivism, which “requires of science that it be independent of an observer” (Williamson “Efficiency” 157). In her article, “Reframing Reliability for Writing Assessment,” O’Neill traces historical uses of reliability, citing Roberta Camp, for example, who identified

some of the key factors that may need to be addressed to develop writing assessments that take into account what we know about writing as well as the principles of fairness, equity, and generalizability—concepts, she explained, that are associated with reliability. (O’Neill “Reliability”)

Fairness and equity are of particular importance to twenty-first century WPAs and assessment practitioners; however, we still operate under a twentieth century “frame” of reliability being limited to agreement between readers. Fairness and equity are prominent in discussions about the consequences of writing assessment and of validity. While there may be potential to reframe the concept of reliability, I would like to turn my attention here to validity, which has been the subject of most scholarship on writing assessment within the writing assessment community.

Validity has been a central feature in writing assessment for teachers and psychometricians alike, inspiring research in both communities for decades. Composition scholarship has consistently provided a brief definition, offered by Edward M. White: “Validity means honesty: the assessment is demonstrably measuring what it claims to measure” (“Apologia” 40). As Peggy O’Neill explains, however, validity is much more

complex. To the educational measurement community, an assessment is valid when its purpose aligns with its consequences. That is, validity is not a *property* of an assessment: it is not somehow located within the measurement. Rather, validity is a result of how the assessment is used or what its function is within a specific context (Messick “Measuring”; Messick “Meaning”). In a sense, validity is a “test” of an assessment, and it is based on what people do with the results.

In our current era, where writing assessment is becoming a much more prominent subfield of composition studies, practitioners and scholars embrace that richer understanding of validity. What is striking here is the shift in how validity has been defined over the past seven decades. With a disproportionate focus on reliability in standardized tests, writing teachers historically have resisted standardized tests of writing ability developed by psychometricians. Throughout the same time period, however, the educational measurement community has consistently offered a more nuanced understanding of *validity* built on the meanings or consequences of tests (see Figure 1), rather than the brief definition offered by writing teacher-scholars (see Figure 2).

Year	Scholar	Definition
1946	Joy Paul Guilford	<p>“In a very general sense, a test is valid for anything with which it correlates” (429).</p> <p>Also noted validity is not a characteristic of a test but is based on a particular purpose (Angoff).</p>
1971	Lee Cronbach	<p>“In particular, what needs to be valid is the meaning or interpretation of the score; as well as any implications for action that this meaning entails” (qtd. in Messick).</p>
1989	Samuel Messick	<p>“Validity is an overall evaluative judgment, founded on empirical evidence and theoretical rationales, of the adequacy and appropriateness of inferences and actions based on test scores” (33).</p>
2006	Michael Kane	<p>“Validation focuses on interpretations, or meanings, and on decisions, which reflect values and consequence. Neither meaning nor values are easily reduced to formulas, literally or figuratively” (18).</p>
2011	Douglas Baldwin	<p>“Validity traditionally was simply defined as meaning that the assessment measures what it is intended to measure; more recently, validity has been defined as being the extent to which the intended meanings and uses of test scores are supported” (328).</p>

Figure 1: Validity Definitions from Educational Measurement Specialists

Year	Scholar	Definition
1994	Edward M. White	“Although validity is a complex issue—colleges offer advanced courses in it—one simple concept lies behind the complexity: honesty. Validity in measurement means that you are measuring what you say you are measuring, not something else, and that you have really thought through the importance of your measurement in considerable detail” (10).
1998	Susanmarie Harrington	“A valid assessment is one which assesses what it sets out to assess (in this case, students’ ability to write in relation to the local curriculum divisions)” (59).
1999	Kathleen Blake Yancey	“Validity means that you are measuring what you intend to measure” (“Historicizing” 487).
2012	Les Perelman	“The term <i>validity</i> in psychological testing refers to the ability of assessment scale or instrument to measure what it claims to be measuring” (“Construct” 121).
2012	Linda Adler-Kassner and Peggy O’Neill	“[Validity] is not inherent in the test but rather refers to the results and their interpretation and use. In other words, an assessment is not valid or invalid but rather produces results that are more or less valid” (74).

Figure 2: Validity Definitions from Composition Scholars

I cite the discrepancy in the concept of validity not to argue for a return to writing assessments being fully in the purview of the “testing experts” with no input from writing teacher-scholars, but to draw our attention to ways the communities might be able to work together on creating better methods of placement and assessment. Closely

examining the language we use and the meanings it conveys is a good place to get started on that work.

The tension between educational measurement and academic communities is often seen as a reflection of the tension between reliability and validity. It is a crucial issue for both writing assessment practitioners and writing program administrators to confront. As Orville Palmer describes it,

The [College] Board regretted the authority of a large and conservative segment of the English teaching profession which sincerely believed that the writing of essays and other free response exercises constituted the only direct means of obtaining evidence as to a student's ability to write and understand his own language. (11)

Some scholars, like Williamson, suggest that a lack of training in psychometrics may have led to misconceptions of classical test theory and opposition to tests designed by testing experts, rather than teachers. At the same time, academic communities faced resistance as they tried to measure and honor the very subjective nature of writing and reading. Compositionists like Ed White struggled to create assessments that represented the field's values and that could withstand scientific critique, as Bob Broad explains:

[the battle in the 1980s] was to legitimize direct writing assessment by making it acceptably 'scientific' within the dominant paradigm of the time. High levels of interrater agreement were critical components of the scientific legitimacy toward which classic test theory aspired. ("Mapping" 261)

A careful review of assessment scholarship uncovers the reality that researchers at

Educational Testing Service (ETS) have initiated some of the most substantial and lasting shifts in writing assessment in pedagogy. Psychometricians, in other words, also demonstrate dedication to the teaching of writing, and it is from psychometrics that the current, complex understandings of validity emerge.

I believe that in the best circumstances, writing assessment is used to improve teaching and learning. Consequently, I propose to shift the conversation in writing assessment from documenting the tension between reliability and validity or educational measurement and academic communities, to instead, focus on its relation to those individuals most heavily invested in and affected by the process. As Brian Huot and Richard Haswell highlight, students are the most vulnerable participants in the assessment cycle. Therefore, an approach to assessment that recognizes students as individuals is essential.

Validity and Placement into Composition

It is useful to begin with the simplified definition of validity – an assessment practice measures what it says it will measure – to initially consider the placement process. If a test states it will measure students' understanding of Edited American English, for example, and presents questions related to grammar, usage, and mechanics, it will likely be considered valid. Research in assessment since the 1970s, however, questions how useful grammar- and usage-based exams are for placement into writing courses (see Chapter Three for an in-depth analysis). The validity of such exams is suspect because it is unclear how well grammar tests represent students' writing abilities.

Placement testing holds a unique position within writing assessment. It evaluates student writing ability and is used to predict how well a student might do in a given

course. Placement is one way we determine students' proficiency, but we use it after they have been admitted to the university. In this sense, it is a sorting tool, not a requisite for admission. It determines which students need the most assistance in their writing, which students are already writing at college levels, and it enables instructors to customize their classes to suit their students' needs and abilities. Placement determines course assignments for instructors while it communicates information about each institution's specific curriculum. Validity, therefore, can only be determined when considering placement as a part of a writing program's entire curriculum and assessment process. Ideally, writing programs would create placement methods that accurately determine students' readiness for college writing in their institution, and then use that information to create curricula, evaluation methods, and professional development opportunities for faculty. In Susanmarie Harrington's words, the placement method would communicate what the program values about writing to its students and teachers, and it would play a central role in determining many of the local decisions WPAs would make. It is essential to consider the writing program as a whole in determining the validity of a placement procedure.

As we reconsider various definitions of validity and pay careful attention to local decisions, we must consider what Sandra Murphy labels cultural validity, and what Asao Inoue and Mya Poe call racial validity. In considering the local consequences of placement decisions, WPAs have ethical and legal obligations to treat diverse populations of students fairly. We must understand those consequences, particularly for students of color and of historically marginalized groups. In working toward a theory of placement, we would do well to recognize the dialogic nature of assessment, to recognize the social

structures it represents, reacts to, or redefines. Placement and writing assessment methods need to reflect how “students make sense of test items and test situations,” as Murphy urges, and should treat students as active participants in the decision-making process (236). This position undoubtedly challenges the usefulness of large-scale standardized exams.

Large-scale, standardized testing has long been critiqued for racial and class bias, and it has been considered to predict little more than a family’s wealth or a student’s success on similar, future exams. As far back as the early 1980s, researchers and citizen advocates, including Ralph Nader, criticized ETS for test bias. As Robert Green and Robert Griffore explain, test bias appears in a number of ways, including unfair or inappropriate content, inappropriate test norms, differential testing situations, inappropriate test use, and students’ varying degrees of test anxiety. In terms of how large-scale tests are created, factors such as who writes the test items, what groups the items are sampled on, what dialect differences exist between creators and test-takers, and irrelevant test items (relying on white middle class common knowledge rather than information contained in the test, for example) all contribute to test bias. Students of color and of lower socioeconomic status are less likely to have access to testing-specific resources, and may be more likely to experience test anxiety or may be more reluctant to guess on answers (Green and Griffore). Peter Sacks, for instance, claims “for every additional \$10,000 in family income, a person gains on average 30 points on the SAT” (Berlak 70), which is evident in research ETS publishes about the SAT:

WILLIAMS: Well, what about the table that shows that people who are children of folks who earn more than a hundred thousand dollars do far better than people who are poor, children of people who are poor?

Ms. RIGOL: I know. I know. And just think for a moment what schools do the people, the children of people who earn a hundred thousand dollars or more—where do they send their children to school? And the ones who earn \$30,000 or less, where do their children go to school? I'm willing to bet that they don't have the same educational opportunities.

WILLIAMS: So—but when you say the test measures aptitude or developed skill with math and with verbal challenges, then you're saying it really measures the quality of the school you've attended.

Ms. RIGOL: You develop these kind of thinking skills, these problem-solving skills, these critical-reasoning skills—you develop these in many different ways over a long period of time. You develop them by reading, you develop them by listening to National Public Radio, you develop them by engaging in dinnertime conversation that is going to challenge you—I mean, many, many different ways. And so these are things that obviously are going to be influenced if you have a very strong academic program. And if you've gone through a very kind of lackluster type of academic program, that's going to be reflected. (Williams)

While the standardized testing industry claims student scores reflect their verbal and mathematical abilities, data on family income suggests otherwise. That is, scores indicate the educational opportunities students have had, typically as a result of their families'

income. Low scores on standardized tests also may reflect students' "resistance to socialization"; that is, it may be a political decision or a struggle for students to maintain their home language and dialect even as they succeed in educational settings (Hull et al. 312).⁷

Testing bias is not limited to large-scale admissions tests, nor is it a thing of the past. As Asao Inoue and Mya Poe explain,

Racism is not about blaming or shaming white people. It is about understanding how unequal or unfair outcomes may be structured into our assessment technologies and the interpretations that we make from their outcomes. This could mean, for example, that a placement exam is racist because it creates a racial hierarchy in a school, placing most African American and Latino/a students into "remedial" courses. In doing so, the exam reproduces social outcomes that arrange groups of people along ostensibly racial lines. (6)

Exams reproducing racial divisions in the university demonstrate the systemic and structural bias Berlak discusses, are unfair for students of color and of low socioeconomic status, and are potentially illegal for writing programs and universities to use. Michael Kane explains that if a work or school policy "has adverse impact on a protected group [via the Civil Rights Act], it has to be counterbalanced by positive consequences [...] if the program is to be considered legally acceptable" (50). In terms of placement decisions, "neither test scores nor other information' should be used to place students in classes where they are 'worse off than they would be in other placements' (Heubert & Hauser,

⁷ There is a wealth of scholarship on code-switching, code-meshing, language, dialect, and identity; see, for instance, the work of Vershawn Ashanti Young, Geneva Smitherman, Elaine Richardson, Rosina Lippi-Green, and A. Suresh Canagarajah.

1999, p. 282)” (Kane 52). Furthermore, in light of research that suggests developmental writing courses may not be beneficial to students and may not help lead them to graduate, we must examine placement practices more carefully (Scott-Clayton; Hughes and Scott-Clayton; Elliot et al.). We have an ethical and legal obligation to treat students fairly in all our educational policies and procedures, especially in placement practices. Writing program administrators and writing assessment practitioners must lead the way here and consider student bodies as we design and review our methods of assessment. As Inoue and Poe claim, considering racial validity may help lead to fairer assessments, including placement procedures.

Writing Assessment in Composition Studies

Writing assessment in composition studies for the past twenty years has been led by Brian Huot, Michael Williamson, and Peggy O’Neill. In fact, Huot’s “principles for a new theory and practice of writing assessment” have become the standard-bearer for good assessments. They have become so central to writing assessment scholarship they are now considered to be “basic principles” (Gallagher “Being There” 450). They stipulate that writing assessment should be: site-based, locally-controlled, context-sensitive, rhetorically-based, and accessible (“Toward” 562; see Figure 3).

SITE-BASED

An assessment for writing is developed in response to a need that occurs at a specific site. Procedures are based upon the resources and concerns of an institution, department, program, or agency, and its administration, faculty, students, or other constituents.

LOCALLY-CONTROLLED

The individual institution or agency is responsible for managing, revising, updating, and validating the assessment procedures that should be carefully reviewed according to clearly outlined goals and guidelines on a regular basis to safeguard the concerns of all those affected by the assessment process.

CONTEXT-SENSITIVE

The procedures should honor the instructional goals and objectives as well as the cultural and social environment of the institution or agency and its students, teachers, and other stakeholders. It is important to establish and maintain the contextual integrity necessary for the authentic reading and writing of textual communication.

RHETORICALLY-BASED

All writing assignments, scoring criteria, writing environments, and reading procedures should adhere to recognizable and supportable rhetorical principles integral to the thoughtful expression and reflective interpretations of texts.

ACCESSIBILITY

All procedures and rationales for the creation of writing assignments, scoring criteria, and reading procedures, as well as samples of student work and rater judgment should be available to those whose work is being evaluated.

Figure 3: Brian Huot's Principles for a New Theory and Practice of Writing Assessment

Huot's principles for a new theory and practice of writing assessment are further described in his 2002 monograph, *(Re)Articulating Writing Assessment for Teaching and Learning*, in addition to his 2009 book, *A Guide to College Writing Assessment*, co-authored with Peggy O'Neill and Cindy Moore. The CCC Position Statement on Writing Assessment, adopted in 2006 and revised in 2009, echoes many of Huot's concerns (see Appendix B). Kathleen Blake Yancey summarizes, "Because literacy is social, this statement claims, assessment must be specific, purposeful, contextual, ethical. And because it is social, we – students, faculty, administrators, legislators – all have rights and responsibilities" ("Historicizing" 500). The Position Statement's various sections state, for example, that assessment should be used "in response to local goals, not external pressure"; should engage "students in meaningful, contextualized writing"; and should use human readers to make judgments.

While most twentieth-century assessment of student writing centered on classical test theory, inter-rater reliability, and indirect or "writing-less" tests, Huot rejects those traditional and modern writing assessments. Instead, Huot builds on the earlier work he and Michael Williamson published in *Validating Holistic Scoring for Writing Assessment* and constructs a theory of writing assessment that draws more on theories of writing studies and writing pedagogy than on educational measurement. Huot calls for deeper connections between the ways writing is taught and evaluated, and for writing experts, scholars, and instructors to be more involved in creating assessments. His goal is to create a new theory of writing assessment that "honors local standards, includes a specific context for both the composing and reading of student writing, and allows for the communal interpretation of written communication" ("Toward" 561). To do so, "We

must also develop procedures with which to document and validate such assessment. These validation procedures must be sensitive to the local and contextual nature of the assessment being done” (“Toward” 561). It is important to examine Huot’s guidelines here to better understand their significance in writing assessment and for placement.

Site-based

Huot’s first principle toward a new theory of writing assessment is that it ought to be site-based; that is, it should be based on and responsive to the concerns of various stakeholders in a specific location. This argument has become familiar in writing programs and within writing assessment, particularly in relation to rubrics, where scholars like Bob Broad and Maja Wilson advocate for teachers of writing working together to develop rubrics that respond to their students and curricula. Rather than using something generalized and created outside of a writing program, Huot and others advocate writing assessments designed in response to and within specific locations. Doing so acknowledges writing as social and as a communicative act embedded in context.

The focus on site-based assessments depends on “the importance of situating assessment methods and rater judgment within a particular rhetorical, linguistic pedagogical context” (Huot (Re)*Articulating* 98). One benefit to a site-based writing assessment is that it evolves naturally out of a writing program, as it responds to a curricular concern, for instance. In this way it can be tied very closely to programmatic assessments, and to observations, interviews, and discussions with the teaching staff and students within a program. That is, site-based practices can embody some of the most generative uses of writing assessment.

While a site-based assessment guarantees attention to the specific writing program and curriculum it is aligned with, it does not guarantee that it will be a good assessment. An assessment may be created on-site, but it still could address concerns expressed by federal guidelines of good writing more than a writing program's. The AAC&U VALUE (Association of American Colleges and Universities Valid Assessment of Learning in Undergraduate Education) rubrics are one such example (see Appendix C). In a recent pilot study to examine writing and critical thinking across the university at UW—Milwaukee, participants were provided with a rubric that favors argument-driven writing and with essays from a FYW curriculum that emphasizes rhetorical analysis. A clear majority of participants in the pilot were from the FYW program, had been teaching in the curriculum between two and twenty years, and understood the kinds of writing students are expected to do in the curriculum. Throughout the scoring and review session, participants described their struggles with their beliefs and values about writing, writing at UWM, and the rubric that required them to award zero points to essays that may have demonstrated critical thinking but had no argument.

The pilot program on critical thinking and writing is just one anecdote. While it clearly demonstrates Broad's assertion that standardized rubrics are typically insufficient to document a writing program's values, it alerts us to the realities in today's political and educational climate. The administrators leading the pilot program at UW—Milwaukee are dedicated to student success and to the nuances of writing across the university; it was clear to all participants that students would not be penalized by the scores they received and that curricula would not be threatened. But there are no safeguards that other administrators would feel similarly. And there are no guarantees that data gathered from

live scoring sessions in the future will only be used for informative purposes. It is very possible the results could be used to reshape writing across the university without input from instructors, faculty, and WPAs. Popular media and scholarship abound with conflicts between university administration and individual programs, and between governmental involvement in writing assessment and educational institutions. It was recently announced, for example, that students' achievement on the Common Core State Standards will be determined in part by Automated Essay Evaluation beginning in 2015 for fourteen states (Barshay). Student and institutional success on the assessments not only provides information about a school or district, it provides avenues to secure or lose governmental funding. It is unfortunately easy to see how "the concerns of an institution...and its administrators" could latch onto public outcries that "Jenny can't write" and attempt to control placement procedures, writing curricula, and assessment practices within a university, ostensibly being "site-based" without being attuned to local programmatic concerns about writing.

Somewhat more in line with those whom Huot envisioned leading site-based assessments are the faculty and instructors who teach, administer, and make decisions about writing on their campuses. To create a system of assessment that is responsive to programmatic needs, however, a site needs to have someone willing to create those assessments. While leaders within writing assessment have called for greater involvement from the rhetoric, composition, and writing studies communities, not every writing program has an assessment specialist on staff. It may also be perceived as too burdensome to create, maintain, and update writing assessments in addition to other teaching, supervising, and administrating duties.

Locally-controlled

Huot's initial explanation of "locally-controlled" writing assessments is familiar to anyone who has participated in programmatic assessment or in designing writing assessments. He calls for individual organizations and programs to control the design, maintenance, and review of assessments. Huot's guidelines highlight the concern in writing assessment scholarship, articulated by William Smith and Peggy O'Neill, that WPAs have often been content with the status quo, or have initiated a new process without considering its deeper theoretical implications *or* being able to consistently review and revise it. As Rich Haswell explains in *Beyond Outcomes*, and Bob Broad states similarly in *What We Really Value*, maintaining an assessment is as important as developing one. It is not enough to create a theoretically and pedagogically sound method in line with course or programmatic goals; rather, it must regularly be reviewed, revised, and updated alongside changes in curricula, policy, and student population. Furthermore, it must also regularly be investigated to ensure it continues to be reliable, valid, fair, and ethical.

For Huot, assessment "should be carefully reviewed [...] on a regular basis to safeguard the concerns of all those affected by the assessment process" ("Toward" 562). The choice of the word *safeguard* here is significant, particularly as it implies that those affected by the assessment process – presumably, students, instructors, and administrators – need *protection*. Jennie Nelson and Diane Kelly-Riley claim, "stakeholders are placed at risk by an evaluation" and "evaluation exposes stakeholders to exploitations, disempowerment, and disenfranchisement" (147). Haswell argues, "never forget that students and their feelings and opinions are the most valuable to explore because they are

the most vulnerable to exploitation” (*Gaining Ground* 189). Some labeling can be helpful for universities and for students. If a student is conditionally admitted, he may be matched to an advisor with more experience with similarly-labeled students. That advisor likely has a lower number of students to work with, whom she can consequently devote more attention to by frequently meeting with students, monitoring their academic progress throughout and at the conclusion of a semester, contacting them by phone, and establishing a relationship with students.

Unfortunately, we often witness exploitations, disempowerment, and disenfranchisement when a placement decision marks a student as “remedial” or “at-risk”; the labels have significant weight in our culture. Glynda Hull, Mike Rose, Kay Losey Fraser, and Marisa Castellano explain “our unexamined cultural biases about difference, our national habits of mind for sorting and labeling individuals who perform poorly, our legacy of racism and class bias” are “the frames of mind” that pervade institutional labeling for students enrolled in developmental courses. And while self-efficacy may be helpful for some students to succeed in college, students whose abilities are undervalued can come to internalize views of their academic deficiencies (Hull et al. 317-18). Remedial labels affect students throughout their academic career, and inform decisions made about their character and abilities: “There is a long, troubling history in American education of perceiving and treating low-achieving children as if they were lesser in character and fundamental ability,” rather than perceiving students as being underprepared or as being capable of doing advanced work (Hull et al. 311). Students do not simply place into developmental writing and proceed with their degree requirements; they are tracked by the university and may be locked into a rigid academic program with

inaccurate assumptions and low expectations. When groups of students regularly fail a course, or are consistently and disproportionately placed into a developmental course, we see the risks students face as a result of writing assessment. Huot's guideline calls for those of us who create methods of evaluation to be aware of the inherent risks and to protect students.

Huot's second imperative makes clear that WPAs, in addition to the individuals, institutions, and agencies in control of creating and maintaining writing assessments, have an explicit responsibility to ensure writing assessments are locally controlled. The Conference on College Composition and Communication reinforces a similar theme, with equally powerful language, in *Writing Assessment: A Position Statement*:

5B. Best assessment practice is continually under review and subject to change by well-informed faculty, administrators, and legislators. Anyone charged with the responsibility of designing an assessment program must be cognizant of the relevant research and must stay abreast of developments in the field. The theory and practice of writing assessment is continually informed by significant publications in professional journals and by presentations at regional and national conferences. The easy availability of this research to practitioners makes ignorance of its content reprehensible.

The Position Statement acknowledges that assessments do not only occur under the direction of individual institutions or agencies; rather, they are affected by, and are subjected to direct influence of, administrators and legislators at local, regional, and national levels. Furthermore, the committee urges practitioners to be as informed as

possible to ensure their methods of assessment are current and serve the best interest of their students and instructors.

Context-sensitive

Just as curricula are typically developed within individual writing programs, the argument goes, so should the methods used to assess local student work also be developed within those programs. O'Neill, Moore, and Huot claim, "Because writing assessment is fundamentally about supporting current theories of language and learning and improving literacy and instruction, it should involve the same kinds of thinking we use every day as scholars and teachers" (59). They cite the varied ways scholar-teachers of writing think about context: in our own writing and analysis,

We examine both the local textual context...as well as the larger social contexts influencing the ways texts are written, distributed, and read. [...]
Context also influences how we design studies and present results. [...]
In much the same way, context informs the decisions we make as teachers.
We consider not only what teaching methods are available but how they coincide with the mission of the school, whether they support the goals of a particular program or course, and how we will modify them for different groups of students. (O'Neill, Moore, and Huot 59)

Huot and his colleagues argue for principles of good writing assessment to be based on principles of writing pedagogy. I believe strategies for twenty-first century writing assessment should reflect the complexities of writing and the transformative possibilities inherent in both the instruction and assessment of writing. Especially since our profession

values writing that reflects students' rhetorical awareness and knowledge, attention to context in writing assessment is vital.

In his homage to William Smith, "Standing on His Shoulders: Understanding William L. Smith's Contributions to Writing Assessment," Huot offers a dialogic perspective on the importance of context with regard to how speakers and writers make meaning. He draws on Michael Halliday's *Language as Social Semiotic* and James Gee's *An Introduction to Discourse Analysis* to argue

language theorists are united in determining that context is crucial for making meaning from any language or literacy event. For Halliday, "All language functions in contexts of situations and is relatable to those contexts" (p. 32). Gee's ideas about context are similar: "The context of an utterance (oral or written) is everything in the material, mental, personal, interactional, social, institutional, cultural and historical situation in which the utterance was made ..." (p. 54). ("Standing" 22; emphasis original)

Huot also reminds us of the role literary theorists like Wolfgang Iser and Stanley Fish have played in understanding writing and reading processes, as they "assert the importance of context and the positions and individual situations that various readers bring with them to make meaning and value out of specific texts" ("Standing" 22).

Understanding the ways in which writing and reading respond to and interact with their contexts is crucial for most disciplines, particularly those under the heading of "English Studies" (and, of course, writing studies and writing pedagogy).

Recognizing the "social life of discourse" is exceedingly helpful to understand the ways we write and read and respond to writing (Bakhtin 259) – and, as Huot contends, it

is helpful to the understanding and creation of writing assessments. In describing assessment systems that rely on instructors reading and discussing student writing, Huot argues, “Without context, readers cannot agree about the relative merit of specific pieces of writing and more important, the decisions to be made on behalf of that writing” (“Standing” 23). A context-sensitive evaluation of student writing is a markedly different perspective from one that emphasizes objectivity. In this approach, the information and experience readers bring to an evaluation of student writing is invaluable; instructors who are intimately involved with the curriculum use their knowledge of a course to make a personal, informed decision about each piece of student writing, each student writer.

Rhetorically-Based

Bob Broad’s work in assessment is founded on the relationship between theories of writing and theories of writing assessment. He claims, “For the assessment to be relevant, valid, and fair [...] it must judge students according to the same skills and values by which they have been taught” (*Value* 11). Following from the argument that principles of good writing pedagogy can guide both instruction *and* assessment, Huot claims writing assessment should also be rhetorically-based: “All writing assignments, scoring criteria, writing environments, and reading procedures should adhere to recognizable and supportable rhetorical principles integral to the thoughtful expression and reflective interpretation of texts” (“Toward” 562). In the third wave of writing assessment, “the endeavors of teaching and assessing writing are theoretically and ethically aligned and are therefore mutually supportive” (Broad *Value* 137). Such positions assume a symbiotic relationship exists between writing and its assessment, that assessment is indeed used to improve teaching and learning conditions. Furthermore, as

Michael Kane discusses, assessments are used to make decisions about students, of course, but also about policies and programs. Writing assessment informs and is informed by its location; “writing assessment here, then, is rhetorical: positioned as a shaper of students and as means of understanding the effects of such shaping” (Yancey “Historicizing” 500). For instance, we create writing assignments that reflect our values about writing – “the thoughtful expression and reflective interpretation of texts” (Huot “Toward” 562). How can we determine whether students have fulfilled the goals of an assignment? By assessing how thoughtful their expression is, how reflective their interpretation is, how well students’ writing matches our expressed values about writing. We do not only create assignments that reflect our values; we also assess what we value (which is evident when different raters give an essay the same score for different reasons, for example).⁸

Accessibility

Huot’s last principle is that writing assessment should be accessible and available to its stakeholders, particularly to students and instructors affected by the evaluations. This appears to be a commonly held and increasingly prominent principle within writing assessment communities, and presents another transformative opportunity for teacher-scholars and WPAs. Increasing students’ access to the principles that guide our pedagogies and assessments increases the transparency of our practices as well as our accountability to students. In my experience, making assessment more accessible has been beneficial to students. At UW—Milwaukee, for instance, the *Students’ Guide* is a

⁸ See Hall, Carla, and Jaffer Sheyholislami. “Using Appraisal Theory to Understand Rater Values: An Examination of Rater Comments on ESL Test Essays.” *Journal of Writing Assessment* 6.1 (2013): n. pag. Web.

central text in each of the three FYW courses. Students read, refer to, and respond to the guide, which includes an introduction to FYW, student learning outcomes for each course, programmatic policies on attendance and grading, suggestions for students to succeed in their writing courses, and resources within the writing program and across the university. The “Goals and Outcomes for Portfolio Assessment” vary for each of the three writing courses, are based on shared principles of rhetorical analysis and reflective writing, and are an integral aspect of the writing program (see Appendix D). Instructors use the student learning outcomes to evaluate student portfolios at the end of each semester, but they feature more prominently throughout the semester: the goals and outcomes help guide lesson plans, assignment criteria, instructor and peer feedback. Students have multiple ways to access the learning outcomes in print and electronically, and often directly work with them in class and in assignments. In my own courses, for example, students and I have explicated targeted learning outcomes before beginning a new writing assignment or embarking on revisions of a draft.

Doug Baldwin suggests in “A Guide to Standardized Writing Assessment” that “students benefit when, with the help of teachers, they become active participants in the process,” a concept Nancy Glazer, also of ETS, addressed in her 2013 CCCC presentation, “Bringing the Test to the Teachers: Building a Bridge to a Standardized Writing Test.” According to Glazer, one way to ensure fairness in large-scale writing assessments is to ensure students have access to information about test formats and scoring. She claims that if creators of exams “lift the veil,” we can provide “a chance to empower teachers, and therefore students, by including them in the process.” Huot and ETS’s dedication to transparency certainly follows writing pedagogy practices wherein

students are more invested when they understand the purpose of an assignment, for example, or the criteria by which they will be evaluated. As a result of these beliefs, ETS has initiated a series of workshops with Historically Black Colleges and Universities (HBCUs) to provide more information about exams like the Praxis Pre-Professional Skills Test in locations without access to test preparation classes or software. The workshops include information about what the exam looks like, what content is covered, where students can find materials to help them prepare, and how the essay is scored. Significantly, to Glazer, the workshops provide teachers with an opportunity to practice scoring essays so they can inform their students about the practice. For many involved in writing assessment, student access to test information results in improved fairness.

(Re)Envisioning Placement

This dissertation follows Huot and Williamson's call to develop a theory of writing assessment drawn from theories of writing that recognize it as "a highly contextualized process that cannot be segmented from the social, linguistic, and situational context in which it occurs" (Williamson "Introduction" 3). In working toward a theory of placement, it is imperative to build on that central principle and consider both writing *and* placement as highly contextualized processes. Recognizing the dialogic nature of assessment lies at the crux of (re)envisioning placement. In Chapters Three and Four, I use rich descriptions of validity and context to analyze current placement practices, which include both indirect and direct assessments of writing, such as multiple-choice tests of grammar and usage, short essays rated by people and by computers, and Directed Self-Placement. In Chapter Five, I articulate a theory of placement, one that takes the history of placement into consideration alongside writers' cultural backgrounds

and experiences. My argument is based on the belief that writing assessment, and placement in particular, can lead to improved teaching and learning; I believe a dialogic perspective can help us fulfill the goal of placement as an “opportunity to learn” (Poe 275). We must situate placement in relation to the various fields, departments, and institutions it affects, to be sure, but we must also view students in our universities as individuals who face very real consequences from placement. Ultimately, a dialogic perspective will enable writing program administrators to (re)envision placement as an articulation of values central to students’ entire university experience.

Chapter Three

The Placement Trifecta: Quick, Cheap, and Accurate

Harvard University's essay entrance exam in the nineteenth century has been well documented throughout composition studies, and has played an important role in discussions about literacy and intelligence, a lack of rigor in high schools leading to students underprepared for higher education, and the birth of basic writing. Lore within composition studies is that the nineteenth-century Harvard entrance exams ushered in an era of viewing students as deficient in writing, an era that increased momentum in the twentieth century with the onslaught of a "literacy crisis" that has plagued the country since (at least) the 1970s. Scholars like Richard Haswell and Robert Connors describe the "horror" concerned parties experienced when half of Harvard's incoming students in 1874 failed the composition exam. Similarly, Brian Huot claims the Harvard entrance exams were created from a deficit model; that is, that students were perceived to be missing essential skills or knowledge for higher education and needed to be tested to determine whether they were prepared for university coursework ((Re)*Articulating* 8-9).

John Brereton explains and challenges the lore within composition studies regarding those famous exams and returns to the primary sources to contextualize the results of Harvard's nineteenth century entrance exams. Significantly, students who failed these exams were admitted to, not rejected from, the prestigious university and were required to take courses to build on their skills. By the early twentieth century, a majority of students were failing the entrance exams at Harvard, Yale, and Princeton, and still being admitted to the prestigious universities:

At each of the Big Three, admission with conditions because a common

pathway to the freshman class; in 1907, 55 percent of those admitted at Harvard had failed to fulfill the entrance requirements. Similarly, at Yale in 1909, the proportion of freshmen admitted with conditions was 57 percent [...] Even Princeton, a smaller institution that was making a vigorous effort to raise its standards under Woodrow Wilson, admitted a clear majority of its students with one or more conditions; between 1906 and 1909, the proportions of students so admitted ranged from a low of 56 percent in 1909 to a high of 65 percent in 1907. (Karabel 22)

A clear majority of students were failing the entrance exams while still being admitted into the universities. According to Brereton, students failed the composition exam, which asked them to pay attention to grammatical and mechanical correctness while summarizing canonical literary texts, at the same rates (or better) they were failing entrance exams in Greek, Latin, and mathematics.

As Brereton explains, the development of “English A,” which has been described as a remedial course,

may have involved recognizing that this was what entering Harvard student writing looked like, and a consequent decision that Harvard had better meet its students’ needs by providing them with the right kind of instruction. The subsequent widespread institution of entrance exams in writing and first year composition courses throughout much of America might well be seen as an accommodation to the kind of student colleges everywhere were getting, a case of colleges adjusting their standards to

reality, although of course blaming the preparatory schools and complaining all the way. (40)

In Brereton's estimation, Harvard's entrance exams – including its writing exams – were useful to help determine which course a student was best prepared for. In this view, the entrance exams were used more for sorting than for gate-keeping. Students were not tested to prevent their admission; rather, they were tested to enroll in courses suited to their skills and needs. The entrance exams could also be used to identify gaps in the university's curriculum and to encourage faculty to develop new courses that accommodated students' instructional needs. Brereton claims educators were “adjusting their standards to reality” in the 1870s, and many teacher-scholars today would agree that placement testing can identify or help create the course that would best meet students' instructional needs.

Reviewing essays written by each student became a daunting, time-consuming task for faculty, however, despite the valuable information it provided. An increasingly diverse student body meant there were no guarantees students had access to a standardized secondary education, access to standardized editions of literary texts, or even that the common texts schools like Harvard tested students on were familiar or available to these new students. The number of students preparing for higher education and high schools exploded at the cusp of the twentieth century, and admissions standards were also shifting from individual institution-based exams on the east coast to verification of high school coursework in the Midwest (Patelis). Harvard's writing exams and subsequent development of new English classes demonstrates, in Brereton's view, how the university recognized what entering students' writing looked like and how

faculty “adjust[ed] their standards to reality” (40). The development of multiple-choice tests to determine writing ability could be argued similarly: universities recognized their shifting, expanding, and increasingly diverse student bodies and used new technology to suit their new needs, their reality. The question that plagues writing teacher-scholars, assessment practitioners, and program administrators in 2014, however, is not necessarily why early twentieth century institutions employed multiple-choice tests. Rather, at the turn of the twenty-first century, the question is why writing programs and universities are still using them, a question this chapter is dedicated to answering.

The Labor of Placement

Placement into First Year Composition is often seen as a necessary evil by compositionists in general and WPAs in particular because it is time consuming, necessary, costly, and inconvenient (Jones “ACCUPLACER” 93). As a form of writing assessment, the practice has several necessary purposes:

to place students in writing courses, to support a WAC [Writing Across the Curriculum] initiative, to support a general-education reform of the undergraduate curriculum, to send a message to prospective employers of degree-holders, to help fulfill a mandate for outcomes assessment from the state higher education coordinating board, and to serve as a form of writing instruction itself. (Haswell “Validation” 129)

Additionally, placement satisfies the needs for a number of participants. As Richard Haswell states, “Institutions use writing placement to recruit students, commercial firms use it to make money, teachers use it to define their courses, students use it to confirm their self-image” (Haswell, “Entrance”). It is crucial to recognize that placement affects

much more than simply the numbers of students enrolled in First-Year Composition courses. As discussed in more detail in Chapter Two, placement supports, represents, and communicates writing efforts within the writing program and across the university, sending messages to students, instructors, parents, corporations, future employers, and governmental bodies.

The commonly-articulated goal of a placement practice is to “effectively [distinguish] the students with the strongest language skills and the students with the weakest language skills from the general population of students” (“Proper Use”). This is where placement gets tricky: because it has such a broad range of stakeholders and participants, it is difficult to determine its purpose, as illustrated in Figure 4.

View of Placement	Purpose of Placement	Stakeholders Typically Aligned with this View
Socially useful experience	Help students succeed academically by determining best-fit writing course.	Educators
Socially responsible test	Sort students into various skill levels for specialized instruction.	Psychometricians, assessment specialists; administrators
Testing as social practice	Communicate writing values to various publics and to students.	Composition teacher-scholars
Testing as obligation	Situates self within the university	Students

Figure 4: Various Purposes of Placement

For educators, placement must be a socially useful experience, one that helps students succeed academically by placing them into appropriate courses. Psychometricians and assessment practitioners are focused on placement as a socially responsible test, ensuring

that it is reliable and valid, accurate and appropriate, fair and ethical (Bennett and Ward). For composition teacher-scholars, placement represents testing as social practice and must therefore be considered in relation to a writing program's curriculum and values. It is not a neutral action; it communicates specific cultural values and expectations to its test-takers and its participants. It can replicate or trouble social structures; it can support or challenge the current era of testing and assessment despair. For students, the most pressing issue is learning which course will be their first in a FYW sequence. That initial placement decision follows a student throughout her entire academic career, determines her time-to-degree, affects her financial obligations, and provides her with an institutional identity.

On the necessary *evil* side of placement is the amount of time it takes to assess students' preparation for college-level writing, determine which course in a FYW sequence would build on students' skills as it adequately challenges them, and re-assign students to different sections or courses because of their schedules, abilities, and desires. If a practice is time-consuming, it is also expensive: writing programs and universities must either hire people or require faculty and staff to do the work of assessment on top of their already-full schedules, which means less time for teaching, advising, service, and research. Members of the FYW administration and instructional staff must schedule time over the summer to determine how and where to place students in the course sequence – both in theory (What are the cutoff scores?) and in practice (Which course should Student X be enrolled in this fall?). Placement is inconvenient, and not only for faculty and staff. Students must schedule their exams, typically for math, foreign languages, and FYW, with a testing center, which may not be on their campus or very near their home. They

must contend with busy spring schedules while still in high school, or summer schedules that often include athletics, jobs, and orientation sessions. Students also have to determine means of transportation, often based around their family's needs and schedules. They can wait months before receiving their test scores, when it is typically too close to the start of a semester to re-test or challenge a placement decision.

Despite these inconveniences, writing programs across the country are convinced that placement is necessary. To mitigate the time and cost of some practices, programs and universities develop and utilize placement methods that are quick and inexpensive. The two most prominent methods that fall under that category were developed in the mid-twentieth century: multiple-choice tests of grammar and usage, and holistically scored essays.

Multiple-Choice Tests of Writing Ability

As more colleges and universities used entrance exams, especially after the G.I. Bill was instituted in 1944 thereby introducing a new social and economic class of students, organizations like Educational Testing Services (ETS) and College Entrance Examination Board (CEEb, later College Board) developed testing methods that could be used nationwide and that were reliable, valid, and cost-effective. These methods included error identification and correction, reorganizing sentences in a paragraph, reading comprehension, and timed short essays. Multiple-choice tests designed to determine students' writing abilities have been scrutinized since their inception, however, and have been resisted by writing teachers for decades. In their landmark study for ETS, Fred Godshalk, Frances Swineford, and William Coffman acknowledge the ongoing criticism that multiple-choice tests of grammar and usage “[proved] nothing about the student’s

ability to write when left to himself,” but the researchers disagreed (v). They “lacked proof,” however, that multiple-choice tests could indeed demonstrate students’ writing ability and embarked on their study to determine the most efficient methods to measure writing ability (Godshalk et al. v).

In 1954, Edith Huddleston summarized her opposition’s resistance, typically represented by writing teachers, to objective tests: they lack the “‘face’ validity” of essay tests, which “require the examinee independently to summon and organize his relevant knowledge. Thus the essay test has been thought of as a ‘natural’ task, allowing a direct approach to important goals” – which critics did not and do not see in objective tests of students’ verbal abilities. Furthermore, such tests have “been severely criticized on the grounds that [they present] the examinee with a task which is artificially oversimplified” by emphasizing “mechanical skills at the expense of style and quality” and by offering a limited representation of students’ abilities because they “may respond correctly in objective tests to items which they do not use correctly in their own expression” (Huddleston 165-66). Huddleston’s explanation is echoed in conversations about the teaching and assessment of writing in the sixty years since her research, and particularly resonates today.

Multiple-choice tests of grammar and usage gained prominence throughout the twentieth century, despite their dubious relationship to students’ writing abilities, for fairly simple reasons:

The multiple-choice question is the mainstay of standardized testing programs in the United States. The format has achieved this position because it permits inexpensive and apparently objective scoring; because

such questions can be answered quickly, allowing broad content coverage within a testing session; and because a sophisticated statistical technology has evolved to support the analysis and interpretation of test results.

(Bennett and Ward ix)

The technology is often built into a standardized test, so the analysis is left to the testing company – which offers additional time- and cost- savings to the university. Campuses that use national, large-scale tests make several assumptions: 1) large-scale testing will yield benefits local tests don't; 2) the vendor will provide research support leading to the establishment of placement scores; 3) the vendor has designed the test based on a traditional view of placement and writing assessment; and 4) the institution will be aligned theoretically and empirically with the purchased test (Elliot et al. 9). They sacrifice local control of placement testing for perceived time and cost efficiency, and expect the vendor to fulfill their local writing program and university's expectations for assessment *or* accept that the test will not demonstrate or promote their values.

Multiple-choice tests are inexpensive: rather than hiring instructors to read student writing or to meet with students and counsel them about writing classes on campus, the scoring mechanism is built into the test. They are also fast: students' answers on computers or Scantron sheets can be verified and provided to the student, writing program, or other interested bodies almost immediately. Quantitative results enable universities and writing programs to compare students' abilities, which leads into determining enrollment in various sections of FYW and supports institutional tracking efforts. The issue of speed in scoring is essential for schools that offer rolling admissions. When students are able to enroll in courses often through the second week of a semester,

the mechanisms for determining their required writing course must provide immediate results. This model does not necessarily account for the scrambling instructors that WPAs experience at the beginning of each semester – which may also indirectly support (or be a result of) the current high numbers of adjunct instructors at many institutions of higher education, who are essentially on-call each semester. That is, a process that allows students to both take a placement test and enroll in courses while a term is in already session can be useful to students, and it can also help a university meets its enrollment and budgetary goals. Finally, if universities use a test students are already taking, like the ACT or SAT, to determine their placement into composition – which the testing companies and assessment literature advise against – the student pays for the exam. This can provide significant cost-savings to a university.

Resistance to the Tests

Edward M. White extends Huddleston's discussion in "The Misuse of Writing Assessment for Political Purposes," as he lists the by now common reasons many writing studies and writing assessment specialists resist multiple choice tests of students' writing abilities, all while being "The least satisfactory method[s] of placement – and the most common in American colleges" (26). He continues: "The multiple-choice test of editing skills does not require the production of text and so measures skills not directly related to the first-year writing course" and these tests may result in "invalid testing, institutional tracking, negative labeling, and retrograde employment practices" (White "Misuse" 26). Additionally, "The indirect relation of such tests to writing is in much dispute and seems particularly weak for students from homes that do not speak the school dialect" (White "Misuse" 26-27). Multiple-choice tests of grammar and usage turn writing into something

that is not personal, that is not an act of interpretation between writer and reader, that can only be done *one* way, that is dissociated from an audience, purpose, and context.

White and other critics of multiple-choice testing question the correlation between editing skills and writing. The proliferation of First-Year Writing courses indicates that writing is a complex process, a process that typically requires a minimum of one semester for students to begin to attain even a minimal level of proficiency. FYW courses certainly emphasize revision, but editing and proofreading are typically left until the end of the semester, when students can spend time polishing their writing. Editing skills and usage identification represent lower-order concerns in a FYW course or sequence focused on students becoming better communicators. Editing at or near the end of a writing process in FYW reflects common professional practices, where writers revise first based on content and polish their prose, for instance, before publication. For White, it is incongruent – and even invalid – to use a test of editing skills to place students into a course where they are expected to do a great deal of writing. It is also discriminatory to use such a test to penalize students who do not use the dominant dialect.

Objectivity and Standardization

Goldshalk et al. indicate the tension between the educational measurement and composition communities, which continues today. Psychometricians historically found the multiple-choice tests eliminated a reader's personal preferences and were more objective than requiring students to produce an essay. Writing instructors resisted multiple-choice tests for a number of reasons: they lack face validity because students are not required to produce writing; they oversimplify writing; they conflate editing with writing; they dissociate writing from its communicative purposes; they are *too* objective

for a personal, subjective activity. Instructors questioned the relationship between students editing and correcting isolated sentences and students communicating effectively in writing. Psychometricians, however, “believed the interlinear and the objective sections of the test were more effective than their critics asserted” and “that the one-hour English Composition Test does an amazingly effective job of ordering students in the same way as a trained group of readers would after reading a sizable sample of their actual writing” (Godshalk et al. v; 21). Tellingly, the focus here is on reliability, which as Peggy O’Neill, among others, has documented dominated the practice of writing assessment for most of the twentieth century (“Reliability”). The task for Godshalk, Swineford, and Coleman – and other psychometricians developing measurements of writing ability – was to demonstrate that the multiple-choice or objective test could sort students into their writing courses *as well as* expert readers could sort students based on a piece of their writing. At this point in the history of writing assessment, the testing mechanism needed to be as good as the human readers were. As I discuss in the following chapter, the advent of Automated Essay Evaluation (AEE) changes that task.

The appearance of objectivity is crucial to understand the proliferation of multiple-choice tests and to understand public discourse on writing. Multiple-choice tests, which test reading comprehension as well as grammar and usage and are created by individual universities, state university systems, and national testing corporations, are viewed as being reliable because only one correct answer exists for each question. Using multiple-choice tests of grammar and usage in place of student writing eliminates dissension. There can be no difference of opinions and ideas about student writing or writing capabilities because the student either demonstrates knowledge of a convention or

does not. For some scholars in the educational measurement community, such tests are also thought to be fair because they are standardized: “all test-takers receive parallel assessments that are given under the same testing conditions and are comparably scored” (Baldwin 331). In addition to providing the same conditions – for example, the same amount of time for each test – test creators work to determine which conditions are most fair. This goes beyond simply determining whether paper- or computer-based tests are fairest for the largest numbers of students. Doug Baldwin of ETS provides a helpful example regarding the use of spell-check on computer-based tests. Creators must determine whether to allow the spell-check feature in writing assessments, and if so, whether students should also then be evaluated on word choice. Furthermore, Baldwin questions how fair it is to permit spell-check when some students may not be accustomed to using the tool – and how fair it is to use either a computer-based test when some students are more comfortable handwriting responses, or a paper-based test when other students are more comfortable composing electronically.

Standardization ensures test creators are attuned to potential biases. Baldwin cites as an example a prompt that required students to write about their favorite music, which clearly excludes students who are deaf/hard of hearing. But standardization does not – and perhaps *cannot* – prevent the widespread critique of testing bias as discussed in Chapter Two. Bias has been cited due to both homogenous test creators and sampled student bodies. Even as test creators and student bodies become more racially, ethnically, linguistically, and economically diverse, standardization cannot account for the differences, skills, and needs present on each local campus or regional university system.

In the face of resistance by writing teachers, Godshalk et al. reject the criticism that a measurement of writing ability should actually measure writing (3). They contend that a student's ability to write well can be broken apart into the ability to recognize separate aspects of writing. I believe, however, that asking students to identify language errors is not the best method to determine their writing ability, particularly in the twenty-first century when many students have not had formal, Edited American English grammar instruction since elementary school. My argument in this dissertation is that indirect measures are at best approximations. Knowing how to edit a sentence does not indicate that students know how to write a sentence, paragraph, or essay. And if we return to examining what our methods of placement communicate to various participants, the multiple-choice test still lacks face validity as it distills writing into context-free tasks.

Validity

Significantly for this dissertation is Godshalk et al.'s assertion that "it must be obvious that a test of composition skill cannot be equally valid in all colleges as a predictor of performance in freshman English" (3). This statement by the psychometricians leads to several important implications. It indicates that one test of "composition skill" may not be enough to determine which writing course will best serve a student. That is, universities might be able to better predict student performance by using multiple methods of assessment, a now-common call in assessment scholarship in both educational measurement and academic communities. For instance, when CEEB was initially skeptical of including an essay in their exams, they found that the combination of objective questions *with* an essay was the most valid method of testing

since it demonstrates students' writing as well as their knowledge of the various aspects of writing (Godshalk et al. 42).

If one test of composition skill is not equally valid in all locations, there must be significant differences among FYW programs at various universities or in different settings. This causes a significant schism in the argument for using large-scale, standardized, multiple-choice exams for placement. Godshalk et al. recognize that it is difficult to use one test in multiple settings, settings that may vary widely. Newly-admitted students at Harvard University typically have different academic skills and needs than newly-admitted students at UW—Milwaukee, which employs an open-access, holistic admissions system, and admits around 90% of its applicants (Brusin). Furthermore, placement tests do not only provide information about a student's writing ability. Writing programs and universities must interpret scores to predict which course would allow the student to draw on her experiences, and challenge her to build on her skills. In this sense, the objective score is much less helpful than *how that score is used* to make a decision about the writing courses available on campus. To use a drastic example, if each student in an entering class scored below the cutoff for an entry-level FYW course, English 101, they would all place into a developmental writing course. The university and writing program would have to consider what it means to write at a “developmental” level in their location, and to examine whether the test aligned with their purposes for it, whether it was valid in their situation.

Finally, the recognition that one test is not valid in all situations points toward the importance of the social consequences of an entrance exam. Just as Godshalk et al.'s claim calls for careful attention to validity, it also calls for increased attention to the

social responsibility of testing. Those who use multiple-choice placement tests must determine how well they correspond to the purposes and intended uses of the test (that is, determine their validity) in addition to how well it provides information about a specific university's incoming students as well as the range of writing courses students may be eligible for. In the midst of a positivist ideology that emphasized writing could be evaluated objectively, in the midst of research that emphasized reliability above – and at times, at the expense of – all else, Godshalk et al. make clear that an assessment must be at least somewhat context-sensitive. Their claim that a test of composition skill cannot equally be valid in all colleges points toward the various ways writing is taught at institutions, the various skills and experiences students bring to universities, the various ways students fulfill their writing requirements and achieve student learning outcomes. Although the composition field's received history paints a conflict between the educational measurement and academic communities, studies like Godshalk et al.'s remind us of the shared concerns about socially responsible testing, and, as Brereton discusses, the importance of adapting FYW courses to students' abilities.

Consequences

Presenting an institution with an assessment method with quantifiable results, high reliability coefficients, and claims of low costs is certainly appealing, especially in this era of decreased funding and slashed budgets. But what are the real costs of such assessments? Some universities that employ placement methods with questionable results allow students to challenge their placement decisions, often up through the first two weeks of a semester. Doing so also increases costs with regards to labor. Instructors may need to diagnose student writing and petition for a change, which WPAs must review

along with student-submitted challenges. Administrators must work to move students between sections, open new sections and close others, and reassign instructors or relieve them of some teaching duties (and for adjuncts, the salary that goes along with them).

Alarmingly, placement decisions based on multiple-choice tests often have negative and long-lasting repercussions on students. While some students may demonstrate such proficiency in Edited American English that they are exempt from a required writing class, they may find themselves behind their peers or struggling in upper-level courses where they are expected to produce a great deal of writing. Students who place into developmental writing courses, in particular, face a number of challenges, as do their faculty. Administrators must struggle to maintain funding for developmental courses, in part because they communicate the wrong message to the concerned public.⁹ Some members of the public resent having to “pay twice” for students enrolled in pre-college courses while at a university, and believe their taxes should be used more effectively for higher education (this argument neglects the decreasing amounts of funding state schools currently receive). Others believe universities should only teach university-level courses, and that there should be no developmental courses on campus (this argument neglects the idea of adapting curricula to students’ needs and experiences). Many are also concerned about the high numbers of students enrolled in developmental coursework because it can indicate institutional discrimination of students from diverse backgrounds outside the dominant dialects, social classes, ethnicities, or races. Still others question the financial burden students bear when they enroll in courses that do not

⁹ See, for example, Nevada’s recent state mandate to decrease the percentage of students enrolled in developmental coursework.

count toward graduation, that offer high numbers of attrition, and that prolong time to degree.

Mass-market exams can also become quite expensive by placing students into courses they are unsuited for or by making the path toward graduation more challenging. Some whose skills are overestimated may struggle in an advanced writing class, which can increase the length of time it takes to graduate. If a student fails a course multiple times, which can happen when she is unprepared for that specific class, it may well affect her financial aid eligibility. Placing a student into a course for which she is overqualified also increases her likelihood of attrition: it requires her to begin her university experience by taking coursework that often does not count toward a degree, and lengthens her time toward that degree. Since many schools use a “lock-step” or required sequence of FYW courses, succeeding in a developmental course means the student must still typically pass two additional writing courses. At UW—Milwaukee, students who begin in developmental writing overwhelmingly succeed in the next course but fail in the final course in the FYW sequence, which also happens to be the General Education Requirement for writing. Students who place into developmental writing therefore can likely face two years of writing instruction before being labeled “college-ready.”

While the front-end costs may seem to be affordable, the practice of using a test designed with other students in mind can be quite expensive. In addition to the costs of labor and for students, these large-scale tests can be expensive because they simply do not provide helpful information for writing programs:

The temptation to accept a commercial test that claims to measure what the campus is looking for is hard to resist, as it is both convenient and

cheap to adopt an existing measure. But no matter how economical such a choice may appear to be, it can become extremely costly when the information it produces is not the information needed. (White *Assigning* 101)

If an exam professes to measure a student's writing ability, but instead measures her editing skills, it is not very useful; in fact, it is invalid. Furthermore, it can be an expensive cost to the institution: if administrators pay for an assessment that doesn't measure what the college or university needs, they have wasted their time and money – neither of which they can afford to do.

Holistic Scoring of Essays for Placement

While multiple-choice tests of grammar and usage are still widely in use nationwide, the field of composition has resisted using them to determine students' writing abilities since their inception. Writing instructors and WPAs have instead argued for placement tests that require student writing. In this way, the argument goes, instructors and administrators can see what students produce and determine which required First-Year Writing course is best for them, and they can simultaneously honor the field's most important values about writing. Such thinking supports their view that the goal of writing assessment should be to determine "a writer's ability to communicate within a particular context and to a specific audience that needs to read this writing as part of a clearly defined communicative event" – which is impossible in "writingless" tests (Huot "Toward" 559). The holistic scoring of essays is also a relatively inexpensive method of comparing students' scores and sorting them into various FYW courses (Cooper 1).

In their introduction to *Automated Essay Evaluation: A Cross-Disciplinary Perspective*, Mark D. Shermis and Jill D. Burstein state, “Skeptics often forget that although we seem to recognize good writing when we see it, we are often at odds when it comes time to articulating why the writing is good” (xiii). This desire to articulate what good writing looks like helped shift writing assessment from mostly multiple-choice testing in the 1950s to multiple-choice and short essay testing, then to essay testing by the 1970s. Brian Huot suggests holistic scoring rapidly gained prominence within writing studies because readers could respond directly to students’ writing:

One reason holistic scoring has gained acceptance so quickly may be that it so well fits this era in English studies. By employing a rater’s full impression of a text without trying to reduce her judgment to a set of recognizable skills, holistic scoring is linked theoretically to recent advances in linguistics, composition research, and poststructuralist literary criticism (White 18) – all of which favor a contextual functional theory of written communication. ((Re)*Articulating* 201)

Holistic scoring records readers’ general impressions of writing. It requires readers to respond to a piece of writing as a whole; that is, rather than evaluating say, a purpose, organization, coherence, and mechanics individually, readers consider the text holistically. Their task is to determine how the writing communicates its message, how each individual aspect contributes to the text as a whole.

Because holistic scoring gained prominence in an era that privileged scientific objectivity, scholars and assessment practitioners in the early days of holistic scoring (the 1970s-80s) focused their efforts on demonstrating the practice’s reliability, often at the

expense of understanding its validity. Research efforts at the time focused on how to achieve consistency between different readers:

Much has been written about how to obtain and measure agreement among raters in holistic evaluations. Often overlooked, however, is the fact that holistic evaluations yield nothing more than relative, impressionistic judgments that cannot give detailed information about writing abilities. (Faigley et al. 2005)

While the field may have been eager to read student writing to determine students' writing *ability*, holistic scoring simply provided readers' quick impressions. Readers are often encouraged to read an essay as quickly as possible – sometimes within 60 seconds! – and to not think too hard about a score (Farley). Holistic scoring can provide writers with a readers' general response to their work, however, reading so quickly typically offers little specific information about writing.

Rubrics

To help readers read more precisely, writing programs and testing companies have developed standards to serve as a holistic rubric. Many composition teacher-scholars have multiple experiences with rubrics, and may be able to relate to an experience I had in an AP scoring session for the English Language and Composition (see Appendix E). Scores could range from 0-9, and we were instructed to first determine whether we were reading a passing essay (the difference between adequately and inadequately developing a position), look at the even numbers' descriptions, and then decide whether the essay fit an even number's description, was a little more effective, or perhaps demonstrated a little less success. The guide seemed to be relatively standard.

In practice, however, I realized there was another, undeclared guide at work. Each description requires readers to examine the writer's position and evidence or explanations of the position. In one instance, I rated an essay an 8, following the scoring guide:

Essays earning a score of 8 effectively develop a position on the relationship between certainty and doubt [the essay prompt]. The evidence and explanations used are appropriate and convincing, and the argument is especially coherent and well developed. The prose demonstrates a consistent ability to control a wide range of the elements of effective writing but is not necessarily flawless.

The essay in question developed a clear position that was especially coherent and demonstrated near-flawless (and powerful) writing. I was pulled aside, however, and was told that the essay was actually a 4, inadequate:

Essays earning a score of 4 inadequately develop a position on the relationship between certainty and doubt. The evidence or explanations used may be inappropriate, insufficient, or less convincing. The argument may have lapses in coherence or be inadequately developed. The prose generally conveys the student's ideas but may be less consistent in controlling the elements of effective writing.

Despite the quality of writing, the creative ideas, the coherent and powerful development of a response to the philosophical relationship between certainty and doubt, my table leader explained that the student was penalized for responding to the prompt with *explanations*, rather than *evidence*. The perception was that examples were concrete, whereas students could make up their explanations. The examples they provided didn't

matter – and were often incorrect, as Les Perelman and Todd Farley discuss – as long as they existed and were not written in the form of “explanations.”

As the scoring continued, it became evident that students were rewarded for writing a five-paragraph essay. In an environment where colleagues reported reading 6000 essays in one week, or about 2.5 essays per minute, speed is crucial. Raters reward essays that follow the traditional five-paragraph structure of an introduction with a clearly stated thesis statement and three pieces of evidence that support the thesis, three body paragraphs dedicated to explaining the evidence to prove the thesis, and a conclusion that restates the thesis and summarizes the evidence. Other raters in my session explained they only read the first page of student writing, citing it as the only way it is possible to “get through” so many essays. Still others, including a fellow graduate student who prides himself on treating students ethically, recommended I forget about any “best practices” of assessment I might consider along with my personal values when I read. I came to realize that this is a common situation that occurs when “the aim of mass-market testers is not education; it is an obese bottom line on the balance sheet” (Perelman “Mass-Market” 435). In fact,

when essay examinations are used for mass testing, they contract most of the drawbacks attributed to objective tests. In the interests of reliability and speed, scorers are provided with a checklist of points to look for. This, along with time pressure, obliges them to score in a mechanical way that is more appropriate for a machine than for a fatigue-prone human. Thus, along with the multiple-choice tests, they do not really answer the question, “What does this student know about X?” Instead they answer the

question, “How many of the following list of Xs does the student know?”

(Bereiter viii)

When people are primarily focused on speed in reading essays, their priority becomes finishing the task quickly or scoring a certain number of essays a minute/hour/day/week, not upholding the field’s best values about writing instruction and assessment. In holistic scoring sessions – whether in local or mass-market situations – the focus on inter-rater reliability often leads to over-simplified scoring criteria. Examples become preferred over explanations one year because most readers can latch onto the idea and score a piece of writing similarly. Ultimately, students are rewarded with high scores – and earned college credit – for knowing the genre and for facilitating the speed-reading raters perform.

Those who demonstrate variety in their writing, or deviation from the pattern, are penalized. I was disappointed with my AP reading experience as an instructor who is dedicated to expanding students’ understanding of the complexities of writing, as a practitioner who believes in the generative power of writing assessment, and as a writer who thought critical thinking and creative expression in writing was valued in higher education.

Reliability

William Smith claims a focus on reliability “does not take into account either the artificial nature of the scales we use or that good raters can legitimately disagree” (“Importance” 300). While psychometricians and compositionists alike have focused on devising methods to ensure agreement as one way to ensure a test is fair, Smith suggests this is not as valuable a task as encouraging real readers to respond to – and discuss their disagreement of – student writing. Smith validates dissent and difference in reading,

responding to, and scoring student writing, which Burstein and Shermis explicitly disparage. To compositionists, deviations from the standards demonstrate the very real readers behind the scoring; the different ways readers respond to writing and the different attributes individuals think contribute to good writing.

In their collection, *Construction Versus Choice in Cognitive Measurement*, Randy Elliot Bennett and William C. Ward of ETS focus on some of the differences between multiple-choice and “constructed response” testing, where test-takers are required to provide a written answer to questions; the length of their response may vary, depending on the testing situation. Proponents of constructed responses argue it can “measure different skills and promote deeper learning than...multiple-choice measures” (Bennett and Ward ix). Detractors explain their objections: it is time-consuming for students to create their own responses, so fewer questions can be asked. Fewer questions limits the breadth of content students can demonstrate. Finally, it is difficult to ensure standardization in constructed-response tests, which leads to a reduced ability to generalize results across student bodies and different institutions (Bennett and Ward x).

While both psychometricians and educators are motivated to achieve more socially useful and socially responsible tests, their means of arriving at the measurements are drastically different. Psychometricians are driven to create reliable, fair, and representative tests, while those concerned with social policy are dedicated to tests that have verisimilitude, that send the right message to those concerned with education, and that can lead to increased success for learners (Bennett and Ward xi). The education community is typically willing to sacrifice standardization, for instance, to arrive at a test with improved social consequences, while that idea is anathema to the psychometric

community. In other words, the division between communities reflects a division between the purposes of assessment: summative or formative. Psychometricians use tests to evaluate student learning, proficiency, or ability, typically comparing students to various standards. Results are used, often in high-stakes situations, to make decisions about student, program, or institution success. Educators, alternatively, use assessments to determine student and programmatic development and use the results to lead to improved teaching and learning conditions.

The Prominence of Writing Assessment in Higher Education

As Norbert Elliot details, conflicts between the educational measurement and academic communities have been at the center of assessment since the early twentieth century. The shift that seems to be happening now, however, is a result of there being for the first time a community of writing teacher-scholars who are also assessment practitioners. While many rhetoric and composition graduate programs continue the tradition of not offering courses in educational measurement theory, growing numbers of writing program administrators and faculty members are taking an increasing role in developing, maintaining, and revising assessment initiatives – and they are now even searching for assessment specialists to join their ranks. There have always been psychometricians concerned with the social consequences of testing, just as there have always been educators concerned with accurate and reliable tests. In the early twenty-first century, as governmental bodies become increasingly focused on post-secondary learning outcomes and results, those responsible for formulating “social policy” must necessarily focus on developing sound measurement tools – there can be no clear division between the two. Socially responsible measurement must consider how to develop reliable and

representative tests *that also* send the right message to constituents and lead to increased success for learners.

Universities and writing programs must consider the cost of any assessment, to be sure, but placement deserves to be given primary and influential consideration, given its vital role in student retention. It is, perhaps, the most public form of assessment on a campus, at least in terms of its consequences for students. Consequently, universities and writing programs must be concerned with multiple “costs,” financial for the university and the student, as well as psychic:

There are certain intellectual activities that standardized assessments can neither document nor promote; these include encouraging students to find their own purposes for reading and writing, encouraging teachers to make informed instructional decisions consistent with the needs of individual students, and encouraging students and teachers to collaborate in developing criteria and standards to evaluate their work. (Moss “Validity”

6)

Standardized assessments – both multiple-choice tests and holistically scored essays – do not permit students to write in ways universities and writing programs value. Multiple-choice tests comprised of grammar and usage *only* test knowledge of conventions. For many scholar-teachers and WPAs, the goal of writing instruction is to help students write more effectively for their individual and social purposes: to succeed in school, to become better professionals, to participate more in society. To achieve these goals, we must build on what students already know, we must “meet students where they are” in the current vernacular, and scaffold their learning to help them become successful rhetors. While

compositionists may disagree about how to enable students to achieve those goals, we agree that we must work to “enable students to function to the limit of their capacity in a society totally dependent on writing, print, and printouts” and in digital environments (Crusius 112). Mass-market, standardized exams, however, completely omit essential aspects of writing, including a demonstration of rhetorical awareness and knowledge; critical thinking, reading, and writing; and evidence of various writing processes, including drafting and collaboration (see the WPA Outcomes Statement, Appendix A). In so doing, they also communicate forms of writing the university and writing program do not promote.

In terms of validity and the social consequences of writing assessment, such placement tests are misaligned with most writing programs. Even in developmental writing courses, knowledge of writing conventions is subservient to rhetorical awareness and critical thinking. In fact, knowledge of conventions is often folded into rhetorical awareness when considering writing for specific audiences or contexts, for example. While some English, Education, or Linguistics majors may take a class on grammar, it is no longer an integral aspect of most FYW courses.

While standardized exams are generally unhelpful in documenting students’ writing abilities, they can be *damaging*. Universities with open-access admissions policies face a significant ideological conflict between the exams and the university mission. Schools implement the policies to ensure diverse student bodies have access to higher education. These “new students” have historically included racial and linguistic minorities, in addition to those who may be socio-economically disadvantaged, first-generation college students, or returning adult students. In urban settings, universities and

colleges express a dedication to representing and improving the metropolitan environment, community, and people. When students are admitted into an institution of higher education, the message is that the student deserves to be there, that they may enroll in college courses and are therefore at the college level in terms of their academic abilities. Returning to the Harvard entrance exams provides helpful insight here:

the eventual establishment of English A does not necessarily mean that it was simply a remedial course.... Establishing the course may have involved recognizing that this is what entering Harvard student writing looked like, and a consequent decision that Harvard had better meet its students' needs by providing them with the right kind of instruction.

(Brereton 40)

I believe that once students are admitted, it becomes the university's responsibility to provide them with instruction that builds on their previous knowledge and facilitates their fulfilling student learning outcomes. Admitting a student communicates that the university believes the student belongs there, is qualified and prepared to be there. When students are also required to take a multiple-choice test of grammar and usage, which often reflects conventions in the dominant dialect, their language varieties and minority dialects are penalized. Such penalties are unethical in an environment dedicated to honoring diversity and access to education, especially when attrition is high for students who are outside the mainstream – and who also are typically required to enroll in developmental courses.

Chapter Four

In Search of Better Placement

The broadest purpose of any assessment is to identify areas that are successful and areas that could be improved. This principle holds true within every day circumstances, such as routine check-ups for a person or for a car, in addition to educational settings, as in writing and programmatic assessment. The history of writing assessment documents a perpetual search to do that “check-up” better, which is especially important for placement into composition. Educators, led by Edward M. White’s work in the California higher education system in the 1970s, fought to incorporate writing into their placement tests, to move away from “writingless” tests and multiple-choice tests of grammar and usage to determine a student’s proficiency as a writer. Godshalk et al.’s validity inquiry describes the situation:

It is generally conceded that objective, machine-scorable questions can be used to measure a student’s mastery of the elements of correct writing, that is, of the rules of grammar, usage, punctuation, and spelling. It is a relatively simple task to develop a series of sentences and to ask the student to identify those which are correct and those which contain errors – particularly if the errors identify the writer as unschooled or careless. But, it is argued, there is more to effective writing than the avoidance of gross errors. One has to be able to organize sentences and paragraphs into an effective whole. The truly effective writer has developed an individual style. He not only avoids errors: he also communicates ideas and moves the reader to feel intensely or to act in relation to the ideas which are

communicated. If the College Board English Composition Test is to be an acceptable test of writing ability, it needs to differentiate beyond the pedestrian level of simple mechanics. (Godshalk et al. 21)

As White and others began to realize some of the problems inherent in holistic scoring of essays, however, they reconsidered how students and educational institutions could be better served by these ubiquitous assessments, and developed new methods to determine students' proficiency upon entering college – methods including portfolios and directed self-placement. That implementation, reflection, and revision is not only a guiding practice within writing assessment, but it is one of the most central beliefs in writing studies and pedagogy. Now, firmly planted in the twenty-first century, we are witness to two camps of placement methods: one in search of cost efficiency, the other in search of cost effectiveness. Directed Self Placement (DSP) and Automated Essay Evaluation (AEE) continue the twentieth century disciplinary tension between teacher-scholars and psychometricians and, in their current applications, could not be much farther apart on the assessment spectrum. But each method demonstrates some of the possibilities that exist within assessment, and pushes us to continually re-envision what placement does and can do for students.

Strengths of Directed Self-Placement

Directed Self-Placement (DSP) was popularized by Daniel J. Royer and Roger Gilles at Grand Valley State University in 1998. While only a small number of institutions use the method – around 30 – DSP has gained prominence within the composition community precisely because it differs so drastically from mass-market tests (CompFAQ). It has been touted as the solution to the complicated process of placing

students adequately and accurately into their FYW courses. DSP is characterized by its attention to local populations, concerns, and curricula, and therefore resists standardization. It also presumes that students can be informed enough to make decisions about the writing courses they are best suited for, and depends on students and faculty to take an active role in the decision-making process. It has two main features: 1) the institution provides each student with information about the writing courses; and 2) each student decides in which writing course to enroll. DSP is recognized as:

empowering students to make choices on their own behalf (Blakesley et al., 2003; Pinter & Sims, 2003; Royer & Gilles, 2003); reducing student resentment about placement (Cornell & Newton, 2003; Royer & Gilles, 2003), communicating with high schools (Frus, 2003); measuring, indirectly, writing anxiety (Blakesley et al., 2003); conceptualizing writing as relatively abstract...orienting students to higher education (Pinter & Sims, 2003); giving students a sense of control over their own writing (Cherneckoff, 2003). (Gere et al. "Assessing" 168)

Directed Self-Placement is purposefully broad in scope so it can better respond to an institution or writing program's local needs, another major difference from standardized exams. In general, students are provided with information about a writing program; its curriculum and sequence of courses; samples of typical course texts and assignments; samples of student writing; and questionnaires to reflect on their previous experiences with reading and writing. DSP relies on students' self-awareness; that is, students must be able to evaluate the likelihood of their fulfilling the expectations, assignments, and outcomes for each course.

Writing programs that use Directed Self-Placement provide incoming students with information about the FYW courses available to them. The level of agency students assert varies across programs: locations may use standardized tests like the ACT to develop suggested guidelines for students, may counsel students individually or in groups, or may assign scores to questionnaires (Inoue et al.). In some locations, students receive an initial placement, often based on their ACT or SAT scores, and then may opt to use DSP to challenge their scores; they may also be required to complete a typical reading and writing assignment as part of their self-assessment (Peckham “Online”). Other locations have students complete a profile and submit a piece of writing read by at least two instructors, who then work with students on their placement decisions (Lewiecki-Wilson, Sommers, Tassoni).

In some of the best-known work about institutions that use the practice, students and faculty both are satisfied with DSP on the whole. They cite cost savings, especially when compared to reviewing portfolios of student writing before making a placement decision, in addition to significantly enhanced interactions among stakeholders in the placement process. DSP has been found to affirm students’ agency since it is built on the belief “that students will be mature enough to choose the course that is right for them, if they have enough information and pressure to choose wisely” (White “Foreword” vii). The majority of studies presented throughout Royer and Gilles’ collection, *Directed Self-Placement: Principles and Practices*, as well as across journals in the field highlight how successful DSP can be. Students value their right to choose a course (Blakesley, Harvey, and Reynolds; Pinter and Sims), were confident in their ability to choose appropriately (Frus; Jones “Self-Placement”), and overwhelmingly believe they made the right choice

(Cherneckoff; Cornell and Newton; Pinter and Sims; Jones “Self-Placement”). Instructors also consistently report positive effects with DSP.

In addition to increasing student satisfaction and participation, DSP changes the ways stakeholders across campus interact in the placement process. It directly involves many university-affiliated stakeholders in the placement process, including students, advisors, administrators, and writing faculty (Blakesley, Harvey, and Reynolds). WPAs and instructors feel like they have increased and improved connections to and interactions with students, and that they can understand their students better. Depending on the method of DSP, this interaction can be even more significant. Instructors may participate in reviewing student self-evaluations, talk with students on the phone, meet with students in person during orientations, or review student materials in conjunction with first-week assignments – all of which heighten their interactions with students and work to develop deeper relationships with students at the beginning of a course. Such interaction can also be helpful for instructors to evaluate students’ work and work habits before the midpoint of a semester, which may increase opportunities for intervention and for intervention to be successful. Higher-level administrators appear to value a writing program’s enhanced interactions with students, as well as the opportunity to offer a personal placement process at a decreased institutional cost (at times). DSP can be used as a recruitment tool for institutions because it underscores efforts to interact with students individually.

It is possible that Directed Self-Placement can increase tensions across campus, depending on pre-existing connections, but the method also has the potential to improve those relationships. In his account of implementing DSP at Southern Illinois University—Carbondale (SIUC), Blakesley describes the difficulties faced as he worked with the

advising center and the Center for Basic Skills, which other WPAs may face as they implement any sort of new program or assessment. Advisors at SIUC, at UW—Milwaukee, and at numerous other institutions, are often the first university representatives new students meet with, and their role is an official interpreter of academic policies and requirements. In terms of placement, particularly in locations with large numbers of incoming students, advisors often receive a score or placement decision and convey that information to students; the placement information does not come from a writing program. Within a DSP framework, advisors shift from relaying information a test provides to providing students with information about the options available to them. This shift requires a substantial amount of preparation from the writing program so advisors are equipped to help students, and so advisors do not emphasize the *directed* aspect of DSP too heavily (by telling a student in which class to enroll) (Blakesley). WPAs with large student bodies must recognize the delicate balance between providing advisors with sufficient information to assist students and enough leeway to recognize their autonomy.

The Center for Basic Skills presented challenges for Blakesley as he worked to implement DSP. Students who were identified as “at-risk” and were provisionally admitted to the university were automatically enrolled in the traditional first-semester writing course. They were not able to select their writing course nor enroll in the pilot stretch program. Blakesley identifies several instances where communication across the university broke down: between the writing program and Center for Basic Skills; between academic and student affairs; between individuals who work with the provost or the chancellor. Blakesley explains the Center’s resistance to SIUC’s new placement program:

I believe now that we were witness to the enactment of an ideology that said students were unable to make good judgments about their writing readiness *because* they weren't good writers. They needed to be led with a firm hand by those who were able to judge them, even if those judgments were unsupported by research and by the record of actual student performance. It is an ideology that transcends our particular circumstances and that Freire so carefully describes in *Pedagogy of the Oppressed*. These Special Admission students, so the argument went, were being saved from themselves and the power would stay firmly in the hands of the ones presuming to save them. The power to place these students, tenuous and unjustified though it might have been, was likely the only form of power that the Center for Basic Skills could wield, or that its director believed existed. (Blakesley 27-28; emphasis original)

While I certainly do not mean to diminish the conflict about DSP at Southern Illinois University—Carbondale, or presume to understand all the factors in play, I do believe this example is useful outside its political context as it represents a common ideological conflict within higher education institutions. It is possible that instead of merely feeling territorial about “their” students, or being convinced that only one space on campus could understand students or know what is best for them, the Center was wary about instituting a new practice with unproven results. The Center for Basic Skills may have seen their students’ success threatened – not only their institutional autonomy – and consequently resisted the program. Rather than obstructing progress, the Center could have been unconvinced that a program that worked at Grand Valley State University could also

work well at Southern Illinois University. GVSU is a public liberal arts school in Grand Rapids, Michigan, that boasted a total enrollment of just under 14,000 students in 1995 (“Grand Valley”); SIUC is a public research and flagship university of the SIU system in Carbondale, Illinois, where enrollment was around 22,500 in 1995 (“SIU”). Grand Valley met with all incoming students in summer orientation sessions to assist with enrollment questions and processes, whereas students enrolled at SIUC in a variety of ways (e.g. off-site, by phone, or one-on-one with advisors) (Blakesley 23).

Blakesley’s discussion indicates that encouraging participation from stakeholders across campus relies on clear communication about shared values and goals, and that buy-in cannot be accomplished in one meeting. Southern Illinois University—Carbondale’s shift to DSP was supported by academic advisors because they were recognized as essential stakeholders in the placement process, were consulted, and were able to continue their advising sessions as they typically had before (with the addition of more information to convey to students). This conflict offers help for WPAs in other situations, and reminds us how vital it is to garner support from all stakeholders across campus to ensure that each participant is recognized for their valuable role, and has an opportunity to guide the implementation of a new program.

Implications of Directed Self-Placement

Programs have decided to use DSP as one way to improve the messages their placement practices convey, to remedy the observation “that real student needs have been grossly out of step with our curricula, pedagogy, and placement practices” (Royer and Gilles 2). Its proponents argue that DSP demonstrates a deep dedication to assessment and its relation to curricula, to students and their ability to make good decisions about

their own learning, to “the importance of self-assessment in writing generally” (Royer and Gilles 8). The practice is not without its skeptics, however, and even those who have used DSP for over a decade, like Anne Ruggles Gere and her graduate students at the University of Michigan, offer advice for using the practice carefully; for example, by providing students with more time than a week or several days before the semester begins to select their course (Gere et al. “Assessing” 164). Gere et al. also suggest programs carefully design DSP to align with the theoretical foundation and curriculum of a FYW program. In analyzing DSP’s first decade at the University of Michigan, Gere et al. explain:

This analysis of various aspects of validity makes visible a number of weaknesses in the form of DSP assessment that existed at the University of Michigan through 2008. Some dimensions of the DSP questions aligned with the construct of writing in FYW, and standardized test scores showed some correlation with responses to DSP questions. However, the central features of FYW were not addressed by the DSP; the time gap between survey completion and course selection diminished the DSP’s substantive validity; scoring of the DSP survey lacked alignment with the construct of writing in FYW; DSP scores did not generalize across time and various populations; the values implicit in the DSP survey differed from those of FYW; and while Practicum had demonstrable benefit in the view of students who took it, many students who might have benefited from Practicum [the course with more instructor support integrated into it] were not led toward it by their experience with DSP. (“Assessing” 170)

DSP is cited as a powerful alternative to standardized exams – as *the* most valid method of placement (Royer and Gilles) – and as an alternative that demonstrates some of the field’s core values about writing, instruction, and assessment. It is therefore essential to carefully review the process because of its potential to shift placement nationwide. Even with an esteemed teacher-scholar at the helm, DSP can face difficulty achieving validity and coherence with FYW programs. Gere et al.’s work suggests that in order for the placement method to fulfill its possibilities, it must preview, represent, and help prepare students for the values and the work they can expect in FYW.

In standardized assessments of writing ability, resistance has existed among both educational measurement and academic communities for at least one century. DSP is presented as an affordable alternative that treats students ethically, that involves the entire campus community, and that ensures public conversations about writing on campus are consistent across curricular and assessment efforts. Early criticism about Directed Self-Placement questions how ethical it is for WPAs to require students to make the decision about their required writing courses when experts from various fields and across decades have been unable to reach consensus about placement. How can students be expected to make a good decision about their college writing course when they do not yet have experience with college writing? Incoming students, according to this perspective, simply do not have enough information about college-level writing or DSP to make reasoned and appropriate decisions (Bedore and Rossen-Knill). Rather than a commentary on students being unintelligent, Pamela Bedore and Deborah Rossen-Knill speak more to students’ inexperience within the university. Many students are simply unaware of the consequences of their enrollment decisions. Their choices have repercussions beyond the

initial semester in which they enroll – that first writing course helps determine their path to graduation, and might not even carry credits that help them along the way. Those who self-select a developmental writing course because they are apprehensive about writing in a university may be locked into a rigid sequence of courses; may be labeled as a remedial student across the university; may face negative self-perceptions as a result of institutional labeling and tracking.

Furthermore, Anne Ruggles Gere, Laura Aull, Timothy Green, and Anne Porter question the relationship between students' past experiences and their future work:

Implicit in the University of Michigan's DSP questions was the value placed on looking to the past. All of the questions were oriented toward what students had already accomplished, both in terms of general achievement and specifically in literacy practices. This was a reasonable stance since past performance is typically considered a good indication of future achievement. However, in looking to the past, the DSP process provided students with little information about the writing that would be expected of them at the University, thereby potentially leaving them with the impression that they would be able to easily match their successes in high school writing in University classes. This was especially problematic because the FYW course looked to the future, aiming to prepare students to write the evidence-based arguments that would be required of them throughout their studies at the University. ("Assessing" 169)

This is the crux of the problem of placement and admissions in higher education. How can we use students' past performances to predict their future success? How can we

appropriately communicate our values, expectations, and goals? Students must consider their previous reading and writing practices, to be sure, but must consider them in a new context. Gere et al. underscore the importance of creating DSP materials that help students understand what writing in their new location entails and echo Ed White in his Foreword to Royer and Gilles' collection about Directed Self-Placement: "DSP depends on the institution clearly defining the requirements and proposed outcomes of its different writing courses, maintaining consistency in those definitions, and then communicating them to entering students" (vii). Publishing specific information about a writing curriculum and its core values can help students understand the work they are expected to do, and it can help communicate a more unified and powerful message to all stakeholders. It can also help to strengthen a writing program by encouraging unity across various sections of writing and writing courses (Cherneckoff).

One way programs can appropriately prepare students to place themselves is to go beyond questions about students' past experiences and move beyond simply providing examples of course texts, prompts, or student writing samples. Irvin Peckham describes the Online Challenge students may complete as a powerful introduction to the writing curriculum at Louisiana State University:

We give students three days to read eight to ten articles on an issue and then another three days to complete the writing task based on these articles. The writing task is in an informative genre: the writers are asked to explain the issue to readers who are not familiar with its details. We want writers to explore the range of positions people take and the different interpretations of data they use to support their positions. We discourage

writers from taking a position on the issue. We ask them to step back from an emotional involvement with the issue and do their best to explain it.

(Peckham “Challenge” 723)

The Online Challenge method offers impressive face validity: What better way is there for students to determine their readiness for a curriculum than diving into it? LSU has also avoided some of the time and cost concerns that exist in conjunction with reading student essays for placement. Their students are initially placed based on ACT or SAT scores and may choose to challenge their score by completing the DSP tasks, which a relatively low number of students, typically those on the cusp of test scores, choose to do.

While DSP relies on students’ careful self-reflection, many students do not have enough experience to thoughtfully and accurately consider what it means to write within a higher education setting:

Self reflection alone would not provide all the information necessary for deciding which writing course to choose. Students would have to know thoroughly the curriculum and assumptions informing each course—a clear impossibility for them. The placement process would thus need to include students working with those that know the courses best—the instructors of those courses. (Lewiecki- Wilson, Sommers, Tassoni 168)

In addition to lacking experience with college-level writing and knowledge of a specific curriculum, students often are unfamiliar with self-evaluation. Ellen Schendel and Peggy O’Neill suggest some students may have internalized the educational gaze; that is, rather than defining their own writing or educational abilities, they focus on what past assessments in school settings have said about them.

For historically marginalized groups, DSP may serve as a gate-keeper, not as a force for empowerment or affirming student agency. By asking about personal and community literacy practices, DSP can emphasize external factors over students' drive to succeed in a higher education setting. By encouraging students to reflect on past educational experiences, DSP can replicate racialized social divisions. If a student – marked by race, class, gender, (dis)ability, ethnicity, culture, language – has always been in developmental courses and DSP asks her to reflect on those experiences, it is likely she will continue to enroll in developmental courses. She may not have had the opportunity to come in contact with the critical thinking, reading, and writing a university expects: the work she has completed in the past might not be indicative of her ability to thrive in the near future demanded by a university.

In an effort to support DSP, Cynthia Cornell and Robert Newton describe their experiences with the practice at DePauw University, a small liberal arts school in the Midwest, and offer insight into how DSP affected different groups of students – European and African American; male and female; first and second generation immigrants. They found that students chose courses based on factors other than their “readiness scores.” Women and African American students typically achieved higher than their readiness score would indicate *and* selected the preparatory course at higher rates than other students. Traditionally-labeled “at-risk” students who chose the mainstream writing course achieved at a higher level than their counterparts who chose the preparatory course. Cornell and Newton conclude that these results indicate the powerful role self-efficacy plays in student achievement, but I want to emphasize a particular concern, namely about how DSP likely affects historically marginalized

students. As Asao Inoue et al. declare, “What is left unaddressed, however, is whether different local social and racial formations perform differently, or are satisfied at different rates, or can be argued to have better placements than other social or racial formations.” African American and female students chose the preparatory course more than their European American and male peers. If we consider self-efficacy, this selection can imply that the students who choose preparatory course lack initiative to fulfill the work of the mainstream course (Ketai). If we consider racial validity, this selection may be a result of internalized racial attitudes replicating discriminatory social structures (Ketai; Inoue and Poe). Moreover, students who enrolled in the mainstream course succeeded at higher levels than their peers in the preparatory course, which Cornell and Newton cite as a result of students’ disposition and drive to work hard – but which could also be a result of the preparatory course being unhelpful for a number of reasons. If the students who enroll in that course are the students who have always been told they belong in developmental writing, perhaps WPAs and instructors should do more to intervene in the DSP process.

Student Satisfaction

Of particular interest to me is the role student satisfaction and, to a lesser extent, instructor satisfaction plays in the DSP scholarship. While many studies cite students’ approval with their own placement decision, or with the ability to make that decision, few discuss *why* that measure of success is so important. The psychic benefits seem clear: relying on students to be active participants in their enrollment for all courses is one way to affirm their agency, and is one way to avoid the resentment or confusion students may experience as a result of a placement decision made for them by anonymous readers or a multiple-choice test. I am interested in further work being conducted regarding the

connection between student satisfaction in their first writing course and long-term success in the FYW sequence or in more advanced writing courses.

Another question related to student satisfaction is: What happens when students *do not* choose the course that is appropriate for them? Gere et al. suggest that in University of Michigan's early experiences with DSP, students who would have benefitted from the preparatory course but did not enroll in it, "were not led toward it by DSP" ("Assessing" 170). Gita DasBender further interrogates this question with multilingual students at Seton Hall University. The study highlights an important, often overlooked group of students in placement studies, and a significant gap in Directed Self-Placement practices. Her study uncovered the result that multilingual students were satisfied with their individual placement decisions; however, their instructors and writing program found the students to struggle in their courses and believed they would have been better served in developmental writing courses. In DasBender's estimation, multilingual students chose mainstream writing courses for varied reasons:

This does not simply reflect eagerness to assert a monolingual academic identity and to assimilate into mainstream courses; it is a revelation of the complex set of assumptions that underlie multilingual students' beliefs, attitudes, and judgments about critical literacy and language proficiency.
(383)

DasBender describes how the institution revised its DSP questionnaire to encourage multilingual students to reflect on their varied literacy abilities, and to encourage the writing program to consider students' linguistic abilities (fluency in writing error-free English) in addition to their self-perception. She concludes that in order to better serve

our diverse body of students, WPAs and instructors need to “be keenly aware of language learning as a long, complicated process for which additional resources need to be in place regardless of the type of course in which students place” (383).

The multilingual students DasBender worked with offer insights into how complicated Directed Self-Placement can be. Even when a student questionnaire or reflection asks all the right questions about their literacy, language, reading, and writing practices, students still enter courses over- or under-prepared for the work of college-level writing. I am encouraged by DasBender’s call to broaden the field’s understanding of language learning and writing pedagogy; while she focuses on multilingual students’ lack of confidence in grammar, for instance, many of us have witnessed the same apprehension in monolingual English speaking male students. DasBender suggests that although some students may select a higher-level course than they might be prepared for, their self-efficacy – along with instructor and program support – can facilitate their success.

Placement through Automated Essay Evaluation (AEE)

The automated scoring of student essays has proliferated in the twenty-first century, and appears to be gaining momentum in light of outcomes-based learning and the subsequent governmental funding tied to institutional success on the assessments; in fact, Patricia Freitag Ericsson and Richard Haswell cautioned in 2006, “The need for a wider audience is urgent because machine scoring programs are making rapid inroads into writing assessment” (2). In 2012, a study presented by Mark Shermis, Dean of the School of Education at the University of Akron, claimed computers rated essays as well as human readers (Perelman “Critique” para. 1). The claim subsequently received a great

deal of publicity, and the finding has been praised in publications like *The New York Times*, *Inside Higher Ed*, and *The New Scientist*. At the same time, educators have opposed Automated Essay Evaluation (AEE),¹⁰ and have attempted to gain public momentum supporting their cause through efforts like a position statement approved by NCTE, the National Council of Teachers of English, entitled, “Machine Scoring Fails the Test” (see Appendix F). The statement directly opposes movement toward AEE inspired by the Common Core State Standards, where “various consortia, private corporations, and testing agencies propose to use computerized assessments of student writing” (“Fails”). In 2013, a group of concerned compositionists, including Richard Haswell, Maja Wilson, William Condon, Bob Broad, Linda Adler-Kassner, Chris Gallagher, Patricia Freitag Ericsson, and Les Perelman, created a petition and website entitled, “Professionals Against Machine Scoring Of Student Essays In High-Stakes Assessment,” which has garnered over 4,000 signatures and coverage in *The New York Times*. Psychometricians and writing assessment practitioners, however, offer several current and long-term benefits to AEE, which deserve a closer look.

Placement is a time-consuming, necessary, costly, and inconvenient measurement, often with questionable validity and reliability. Institutions of higher education therefore turn to AEE in an effort to place students accurately and inexpensively. In most circumstances, students spend 25-40 minutes writing a short essay on a topic they have no prior knowledge of. They are typically expected to read a statement, take a position, and explain their position. The prompt and writing requirement resembles what students have experienced in the holistic scoring of essays in institutions and on large-scale, high-

¹⁰ AEE is a broader term than its earlier iteration, Automated Essay Scoring (AES). See Shermis and Burstein, *Handbook*, for a full description of the language shift.

stakes exams (including the SAT, GRE, and LSAT) since the 1970s. The ACT offers the following information for students to review before taking the exam:

Prompts used for the ACT Writing Test:

- describe an issue relevant to high school students
- ask examinees to write about their perspective on the issue

As a starting place, two different perspectives on the issue will be provided. Examinees may choose to support one of these perspectives or to develop a response based on their own perspective.

Sample Prompt

Educators debate extending high school to five years because of increasing demands on students from employers and colleges to participate in extracurricular activities and community service in addition to having high grades. Some educators support extending high school to five years because they think students need more time to achieve all that is expected of them. Other educators do not support extending high school to five years because they think students would lose interest in school and attendance would drop in the fifth year. In your opinion, should high school be extended to five years?

In your essay, take a position on this question. You may write about either one of the two points of view given, or you may present a different point

of view on this question. Use specific reasons and examples to support your position.¹¹

Prompts for the GRE Writing are similar:

Sample Issue Task

As people rely more and more on technology to solve problems, the ability of humans to think for themselves will surely deteriorate.

Discuss the extent to which you agree or disagree with the statement and explain your reasoning for the position you take. In developing and supporting your position, you should consider ways in which the statement might or might not hold true and explain how these considerations shape your position.¹²

Sample Argument Task

In surveys Mason City residents rank water sports (swimming, boating and fishing) among their favorite recreational activities. The Mason River flowing through the city is rarely used for these pursuits, however, and the city park department devotes little of its budget to maintaining riverside recreational facilities. For years there have been complaints from residents about the quality of the river's water and the river's smell. In response, the state has recently announced plans to clean up Mason River. Use of the

¹¹ Retrieved from: <http://www.actstudent.org/writing/sample/>

¹² Retrieved from: http://www.ets.org/gre/revised_general/prepare/analytical_writing/issue/sample_task

river for water sports is therefore sure to increase. The city government should for that reason devote more money in this year's budget to riverside recreational facilities.

Write a response in which you examine the stated and/or unstated assumptions of the argument. Be sure to explain how the argument depends on the assumptions and what the implications are if the assumptions prove unwarranted.¹³

The essays are then rated similarly to human-scored short essays on mass market exams.

AEE scores essays based on

grammaticality and discourse structure [...] sentence errors such as fragments, and other errors such as with subject-verb agreement and commonly confused word usage [...] evaluations of the quality of a thesis statement, or relationships between two discourse elements, such as the thesis and conclusion. (Shermis and Burstein *Automated* xiv-xv)

Proponents advocate for the method because it distills writing to its most objective features. Overlooking the omission of the rhetorical aspects of writing for now, AEE measures many of the same features traditional methods of placement do. Chaitanya Ramineni of ETS explains, “placement testing for writing has typically and appropriately focused on a narrow range of writing skills, such as knowledge of conventions, essential for succeeding in this larger domain [of writing]” (42). It is presumed that knowledge of

¹³ Retrieved from:
http://www.ets.org/gre/revised_general/prepare/analytical_writing/argument/sample_task

language and organization conventions, for example, can be objectively and efficiently extrapolated to writing ability.

Implications of AEE

As Carl Bereiter explains in his foreword to Mark D. Shermis and Jill C. Burstein's collection, *Automated Essay Scoring*, the first phase of implementing AEE in the 1980s-1990s dealt with proving the method could score student essays *as well as* human raters. In the beginning of the twenty-first century, however, AEE may now be able to score *better* than human raters:

Human essay scorers are not perfect; if they were, it would be a first in the history of civilization. As human beings who have lives outside essay scoring, they are susceptible to quirks and biases carried over from their other lives. They are also susceptible to halo effects; the tendency, when something creates a generally favorable impression, to rate it highly on all counts. The correlation between ratings on style and content is probably a good deal higher than it deserves to be. Computer scoring ought to be able to overcome these human foibles. (Bereiter vii)

For Bereiter, the strength of AEE is in its reliability and in its ability to remove the human aspects of scoring – their quirks and biases – from the process. For many in the educational community, however, those human responses to writing are essential to the communicative process. Furthermore, in academia, professional, and personal spaces, writing is most successful when it has a carefully crafted style and accurate (or at least persuasive) content – those are precisely the aspects of writing that speak to our humanness and that generate responses from our readers. The emphasis on accuracy in

grammar and usage, however, results in an invalid placement method, as perceived by the composition studies community:

Validity research is most efficiently guided by the testing of “plausible rival hypothesis” which suggest credible alternative explanations or meanings for the test score that are challenged and refuted by the evidence collected. Does the test require capabilities that are irrelevant to the intended interpretation? Does it result in performances that under-represent or provide a narrower interpretation than what is intended?

(Messick qtd. in Moss “Testing” 114)

When a placement test privileges grammatical accuracy while the curriculum privileges rhetorical awareness, the test appears to focus on irrelevant content that leads to a narrow interpretation of both the act of writing and students’ writing abilities.

Not all educators are convinced that AEE is a viable option. Norbert Elliot and David Williamson explain the tension:

Because writing is an activity that is so deeply human, its association with formulation is double edged. Because students are encouraged to write fluently or to achieve competency in their knowledge of conventions, a certain degree of formulation is necessary. But when these formulations are used by machines as the basis for assessing writing beyond fluency or knowledge of grammar (Attali & Powers, 2008) there is an inherent suspicion that technology can corrupt the essence of a fundamentally human activity (Ericsson & Haswell, 2006; Herrington & Moran, 2012).

(1)

Placement testing in particular represents many of the tensions in the twenty-first century American education system, including the rising cost of higher education tied to decreased governmental funding, increased class sizes, and heightened attention to writing in all majors and industries. Mark Shermis and Jill Burstein's recent collection, *Handbook of Automated Essay Evaluation: Current Applications and New Directions*, explains,

Despite the fact that machines do not read and evaluate essays using the same cognitive abilities as humans, the concerns about AEE might in part be allayed by the fact that different systems have been shown to achieve similar score prediction outcomes (Shermis and Hamner, 2012). (Shermis, Burstein, and Bursky 3)

The testing community suggests teacher-scholars might be more inclined to use or even trust AEE because the software scores writing similar to how people score writing. From my perspective, however, educators are more concerned about what AEE can evaluate and what messages it communicates. While AEE can place students into writing courses relatively well, the academic community is concerned about how supporting the practice can communicate or contradict their core values about the very human act of writing.

Examining how AEE creators and composition scholar-teachers consider *discourse* offers insight on the conflict. In his *Discourse: A Critique & Synthesis of Major Theories*, Timothy W. Crusius examines composition theories posited by four of the most influential scholars of Composition at the time – James Moffett, James Kinneavy, James Britton, and Frank D'Angelo – in an effort to synthesize their ideas into one coherent theory “incorporating the best of all four” (3). For Crusius, “a theory of

discourse responsive to the experiences of composing must express somehow the complex of interactive motives in the writing process and of interrelated structures in the finished product” (114). Discourse in this sense is related to the rhetorical elements of writing: how writers communicate their ideas, how those ideas are represented throughout a piece of writing, and how audiences interact with those ideas. Shermis and Burstein use the term “discourse elements” to refer to aspects of an essay, such as a thesis or conclusion. It is possible that discourse elements could be related to the interrelated structures in the kind of finished product that Crusius describes; however, a closer look at Burstein’s work in particular offers a more limited view.

In discussing the “many factors that contribute to overall improvement of developing writers,” Jill Burstein and Daniel Marcu include

refined sentence structure, a variety of appropriate word usage, and strong organizational structure. Of course, mastery of the closed-capacities (grammar- and mechanics-related factors) is required if one is to be a competent writer. (209)

Burstein’s chapter in *Writing Assessment in the 21st Century* lists grammar and word choice, mechanics, elements of essay structure (e.g., thesis statement), style, word length, and word frequency as characteristics of good writing and elements of writing that Automated Essay Evaluation algorithms can measure. Returning to the “WPA Outcomes Statement for First-Year Composition,” however, we see a different understanding of what makes for good writing: rhetorical knowledge; critical thinking, reading, and writing; [composing] processes; knowledge of conventions; and composing in electronic environments. “Knowledge of Conventions” states,

By the end of first year composition, students should

- Learn common formats for different kinds of texts
- Develop knowledge of genre conventions ranging from structure and paragraphing to tone and mechanics
- Practice appropriate means of documenting their work
- Control such surface features as syntax, grammar, punctuation, and spelling

Significantly, sentence structure, word usage, organization, grammar, and mechanics are all considered to be indicators of a students' knowledge of the conventions appropriate for their rhetorical situations and genre expectations. They are not isolated features of writing, nor are they the most prominent features of writing. The distinction here makes sense: after all, it is much easier to create an algorithm that determines whether a student has articulated a thesis statement than one that recognizes the “complex of interactive motives in the writing process.” Significantly, proponents of AEE argue that holistic scoring of essays reduces humans to machines checking off a list of features; unfortunately, the machine-based system performs the same function.

The tension about AEE reflects the ongoing conflict between academic and educational measurement communities:

Teachers typically want to understand how the technology works, and whether or not it will address relevant issues that will improve their students [*sic*] writing. Researchers in educational measurement typically have questions about the reliability of the technology. Our colleagues in computer science are interested in the various computing methods used to

develop capabilities for Automated Essay Evaluation and evaluation tools.

(Shermis and Burstein *Handbook xi*)

Shermis and Burstein explain, “A primary challenge is to develop Automated Essay Evaluation and evaluation capabilities so that they are consistent with the needs of educators and their students” (*Handbook xi*). Shermis and Burstein assume that AEE technology is already in place and teachers should understand how it can benefit their students, most typically through reducing the amount of time they need to spend on providing feedback on grammar and usage.

Notably, proponents of AEE cite the obvious up-front cost benefits of the practice in general and for placement tests, and quickly move to underscore the emerging possibilities of integrating AEE into the writing classroom, “of turning AEE into a learning tool” (Bereiter viii). This rhetorical shift, from placement to curriculum, is important to note: teachers are encouraged to use the accompanying classroom software “so their students can pass standardized tests that will be graded, at least in part, by assessment machines” (Ericsson and Haswell 2-3). Rather than spending time grading language, usage, and grammar, teachers will be able to focus on the rhetorically nuanced aspects of writing the concerned public is interested in. That is, “Providing students with instant feedback about grammar, punctuation, word choice and sentence structure will lead to more writing assignments... and allow teachers to focus on higher-order skills” (Stross). The possibilities that can emerge when curricular space is devoted to content and advanced writing features drive much composition pedagogy.

A sinister view of AEE shows the corporations and proponents essentially turning the testing market in their favor: companies create standardized, mass-market tests that

are inexpensive for schools to implement. Teachers then must teach to the test, and companies create software to help them do so. Finally, the companies lobby to keep this cycle in place and make it stronger, for example, by participating in No Child Left Behind and the Common Core State Standards, where the writing students at all levels may be evaluated by AEE. Members of the composition and writing assessment communities are rightfully concerned by the inroads Automated Essay Evaluation is making in both K-12 and post-secondary education. Despite claims that AEE can determine students' writing proficiency quickly and inexpensively *and* can be used as a learning tool to improve students' writing, many members of the academic community are not convinced.

Anne Herrington and Charles Moran highlight the message AEE sends to students: that “humans are unreliable, quirky, expensive, and finally irrelevant; and students' writing matters only in a very narrow range: its length, its vocabulary, its correctness,” or its ability to conform to what a computer can measure” (“What Happens” 497). I believe using such methods to evaluate students' writing abilities or to determine which class a student should begin her academic writing career *undermines* our deepest values about writing and the teaching of writing. The fields of rhetoric, composition, and writing studies have developed because we are convinced writing is of central importance to the university and to our communities. Herrington and Moran claim “writing to a machine is not writing at all” because writing is – at its best – rhetorically situated and responsive to writers' concerns and interests. This view is not necessarily at odds with the amounts of writing twenty-first century denizens do *on* machines because that is most often done with an audience in mind. On Facebook, Twitter, Instagram, and Tumblr, that

audience likely includes friends, family, and strangers, but a writer is able to make informed decisions about her audience. It is much more difficult to make decisions about what a machine or software expects from writing. Even when students are writing within the artificial constraints of the classroom, they are still writing for people, including their instructors and peers. We value critical thinking, for example, evidence of students grappling with ideas and texts and expressing their thoughts in writing. Good writing might include multi-syllabic words and use a variety of sentence lengths, but those are not the qualities that make it good.

Most computer-based placement tests emphasize conventions necessarily: they can only test what a computer can measure. AEE software cannot, for example, detect or evaluate puns:

Consider the following statement: *“I forgot what a boomerang was, but then it came back to me.”* While NLP [Natural Language Processing] methods can be used to provide linguistic analyses of texts, a system would not “get” the pun here, but a human reader would. NLP methods are currently limited to the more literal linguistic analyses. For instance, co-reference resolution tools can be used to determine if “me” in the sentence above refers back to the “I”, and if “it” refers to the “boomerang.” A syntactic parser could be used to identify the sentence as a “compound sentence” with two independent clauses. [...] (Shermis, Burstein, and Bursky 2)

AEE software cannot replicate the “complex cognitive processes for reading” humans use, although it can provide detailed feedback about the linguistic features of a writing

sample (Shermis, Burstein, and Bursky 2). Unfortunately, when decisions about student readiness for a required writing course are delegated to a computer, we send a message to our various publics that we don't think very highly of writing. Students who demonstrate mastery of the conventions of Edited American English by correcting sentences, for example, can earn General Education Requirement credits and "place out" of what could be the only writing course they take at some universities. Those who do poorly are destined for at least three semesters of writing. I question what these actions communicate about writing to the various participants within placement and I am dedicated to uncovering strategies that better align writing assessment with writing pedagogy.

As Paul Deane of Educational Testing Service (ETS) explains, the tension about various assessment methods lies in expectations different communities have about writing. While some psychometricians are most concerned with standardization and reliability, "in a humanistic context...reasoning skills, writing processes, genre practices, and the cultural and social contexts in which genres develop take center stage" (Deane 9). Proponents of AEE have already successfully convinced a large number of institutions that the upfront cost savings are enough to sacrifice our highest values about writing because students are placed "good enough" into required FYW courses, despite any potential negative social consequences of the test. However, as Mya Poe suggests,

Large-scale standardized testing [...] has significant effects on the teaching and learning of writing; the economies of such assessment often result in tests that suffer from construct underrepresentation, which in turn

directs writing instruction away from pedagogy informed by nuanced theories of writing development. (274)

When high-stakes decisions about students are determined by large-scale standardized tests in any form, the teaching of and learning about writing is threatened. If AEE software is introduced into curricula – which is beginning to become more common in high schools and community colleges – the construct of writing will become even more distorted. Students will become preoccupied with “correct” writing, rather than writing that develops an individual style, that displays their content knowledge, that demonstrates rhetorical awareness.

Additionally, Automated Essay Evaluation reduces opportunities for professional development and community building. Reading students’ writing is a time-consuming and sometimes argument-inducing experience, to be sure. It is also, however, a valuable opportunity for instructors and administrators to see how students communicate, how they demonstrate their understanding of current events and difficult ideologies, how they grapple with the big ideas our society considers important. Reading students’ writing also strengthens writing programs. Faculty and staff must articulate what they find valuable or troubling in the writing they read; they share strategies for working through difficult concepts; they challenge others and have their beliefs challenged by the simple act of reading student work together. Relegating the process of assessing student writing to a computer diminishes the opportunities for educators to develop their pedagogies and philosophies, and their familiarity with their students.

Conclusions

Returning to a dialogic perspective of placement can be helpful with regards to the apprehension about AEE. Building on an understanding of the historic tension between academic and educational measurement communities, in addition to the needs of students enrolled in college-level writing courses and the concerns of our various publics, it is possible to uncover benefits of the new technology. Andrew Klobucar, Paul Deane, Norbert Elliot, Chaitanya Ramineni, Perry Deess, and Alex Rudniy – members of academic and psychometric institutions – offer insight in their chapter, “Automated Essay Evaluation and the Search for Valid Writing Assessment.” It has already become clear that AEE can offer insight about students’ fluency in writing, and that the length of writing students submit often determines their placement, the rational being that the length relates to students’ fluency in Edited American English in addition to their experience writing in general and on essay tests (Jones “ACCUPLACER”). While AEE currently offers information about writers’ organization, clarity, relevance, style, wording, and sentence variety, Klobucar et al. suggest that it might not always *only* focus on these features. That is, as the technology has developed, it might continue to do so, and might be able to offer even more sophisticated information about writing that could be useful to students who are learning about and building their skills in college-level writing.

Carl Whithaus takes a similar position, and urges: “Instead of joining the popular, and all too-easy-to-tell, narrative condemning robo-graders, writing teachers, researchers, and program administrators should engage with this research about AEE” (vii). We are called upon to bridge the educational measurement/academic community divide and to think more critically about our resistance to AEE. While previous interactions have been

defined by argument, Whithaus suggests the time is ripe for more dialogue across writing assessment communities:

Does AEE have a place in the future of education, in the future of writing instruction and writing evaluation? The answer is an emphatic “Yes.” And it is an emphatic yes whether or not English teachers immediately embrace software as one of the tools available for generating evaluative feedback. Software will, already is, in fact, reading and responding to students’ writing. Students are using software as an evaluative, responsive tool when they type in Google Docs, on Tumblr, or on their status updates in Facebook. In these social media applications, software is evaluating and providing feedback on the intimate level of words and sentences for writers. For many teachers and students in the next ten years, AEE may become a far more common tool for providing feedback on the paragraph or even essay level. If so, writing teachers and educational administrators need to better understand AEE software. (viii)

Regardless of educators’ resistance to the technology, Whithaus believes AEE will become more prominent in classrooms and educational settings nationwide. Attempting to understand the software, to examine it, to engage in discussions about it will consequently be required of educators. Doing so will also be able to enrich the current conversations about AEE by including informed analysis by expert teacher-scholars. While many, especially Les Perelman, are already doing that work, Whithaus seems to call on more educators to participate in the discussion. I am, however, skeptical about the inroads of AEE into writing classrooms. I can see a clear benefit in some ways: students

can submit their writing and receive thorough feedback on their grammar and usage. Such feedback is appealing to me as a writer who, in full disclosure, would appreciate a greater understanding of grammar rules and conventions in English. I join the chorus of the concerned, however, because I can see how AEE can easily be transitioned from a classroom tool to the dominant form of feedback – which would also be an *efficient* way to increase the number of students enrolled in writing courses.

In his discussion of the “worship of efficiency” in the nineteenth-twentieth century positivist era, Michael Williamson explains how multiple-choice tests of language and grammar were perceived:

[The testing method] was also considered a fairer approach to assessment because it led to more reliable scores as it removed the causes of certain types of errors not associated with the actual performance of the examinee. In particular, it removed the apparent error involved with teachers’ differential evaluations of the same students’ performance, by reducing the decision about whether the examinee had given a right or a wrong answer to a clerical decision, thus replacing teacher judgment with the predictable answers preselected by the psychometric scientist responsible for test construction. (“Efficiency” 157)

Clearly, both Directed Self-Placement and Automated Essay Evaluation “replace teacher judgment.” DSP does so by leaving the decision about a writing course in the hands of the student with varying levels of input and guidance from teachers; AEE does so by training a computer system to read like a teacher who is primarily focused grammar and mechanics. DSP emphasizes core values about writing held by teacher-scholars in

rhetoric, composition, and writing studies, and attempts to shift placement away from an era of efficiency. It is used in such a small number of institutions, however, and presents perceived obstacles about cost and time, that it is not seen as a viable option nationwide. The methods still in use across most universities and colleges are based on replacing teacher judgment with technologies that can score preselected answers or short written responses very quickly. Some might argue that DSP *also* replaces teacher judgment with decisions of students who have no experience or expertise in college-level writing, a critique that gives me pause.

Analyzing the four most prevalent placement practices, DSP and AEE in particular, highlights several tensions I am not quite sure how to address. Requiring students to participate in placement underscores the importance of writing in the university. Using practices that sacrifice the complexities of written expression for efficiency, however, undervalues the importance of writing. Placement by AEE shows a willingness to interact with emergent technologies and to align digital pedagogies with digital assessments (Klobucar et al.) while it also suggests computers are better able than teachers to make decisions about writing *and* that writing is best when it follows simplified conventions. And if that is true, then why do we need teachers of writing? Perhaps because I have been trained in the study of Rhetoric and Composition, and North American and British Literature before that, I am convinced we do need teachers of writing, and practices that demonstrate our deepest-held beliefs. It is not enough to employ innovative curricula and programmatic assessments; WPAs should also work to offer placement practices that align with core writing values. Writing programs that communicate transparency over opacity, cohesion over chaos can be better positioned to

provide opportunities for learning for all of us involved in the teaching and learning and research and assessment of writing.

Chapter Five

(Re)Envisioning Placement in the Twenty-first Century

I began this dissertation believing I could develop a method of placement that would hit the trifecta of being quick, cheap, and accurate; that would adequately communicate the core values about writing teacher-scholars hold; that would draw from and demonstrate the best practices in writing assessment. I end it believing it is difficult to create a meaningful and ethical method of placement in an era of constraints imposed by the wide range of stakeholders – difficult, but not impossible – and that there may well be no one best method of placement for all institutions to use. My hope is that the following suggestions and considerations can help writing program administrators, writing assessment practitioners, and writing studies teacher-scholars (re)envision placement as an essential aspect of a writing curriculum situated within specific institutions.

An Efficiency Model of Placement

The ultimate goal of placement, as a form of writing assessment, is to improve the teaching and learning that takes place within universities and colleges nationwide. Its historical use has been twofold: 1) to identify which students need the most instructional support, often considered to be those who would not earn a C or above in a traditional first semester writing course; and 2) to sort students into various ability levels to help teachers target their instruction and help students be appropriately challenged; that is, to ensure they receive instruction that will build on their current skills and encourage them to develop into college-level critical thinkers, readers, and writers. Anthony Petruzzi explains that the current situation of writing assessment is based on “neo-empiricism”:

Educational assessment, following the human science of psychology, uses standardized writing tests because scientific reasoning assumes it can indirectly observe and measure the transcendental function of writing ability. ... Writing assessment is dominated by neo-empiricism because its methodology legitimates evaluation by producing quantitative data that is reliable and valid. Both direct writing assessments and holistic scoring accommodate the scientific methodology of measurement. (219)

This model stems from a twentieth century perspective that considered writing to be comprised of a quantifiable set of skills, where students were sorted based on their verbal abilities because of their relationship to written abilities (Huddleston). It is reminiscent of an assembly line: students who demonstrate traits 10-15 are sorted into the English 102 line, students who demonstrate traits 5-9 are sorted into the English 101 line, and students who demonstrate traits 1-4 – or unidentifiable traits – are sorted into the developmental English line.

The historical view of placement is focused on efficiency in the instruction of writing. Placement has been seen as necessary to increase the productivity of both instructors and students in writing classes. Some institutions, for instance, segregate students who score highest on placement tests or entrance exams into honors level courses where they receive more advanced instruction than in a typical FYW course and where they are surrounded by their exceptional peers. Students sorted in this way are alleviated of the burden of assisting their peers who may have less preparation or poorer writing abilities, and instructors are rewarded with teaching the best prepared and most motivated students. Mainstream FYW courses are then filled with students who are

typically average, and developmental courses with students who need the most instruction, so teachers can target their lessons, assignments, and assistance appropriately for each group of students. While a similar “division of labor” sorting process does not exist consistently across institutions or even within departments, it is perceived as being necessary to efficiently shuttle students through their required writing courses. Many FYW courses nationwide face increasing cap limits (often above the 15-20 student per writing course recommendation made by NCTE) in addition to increasing numbers of untrained graduate teaching assistants at the helm of the courses; within this setting, sorting based on abilities can help ease some of the labor of teaching.

An efficiency model of writing and placement treats FYW as necessary but burdensome for both students and the university. Writing program administrators and composition teacher-scholars, however, have made great headway in shifting the conversation about FYW from teaching students basic skills in a content-less discipline to instilling core values and perspectives on writing effectively within a range of discourse communities. In fact, effective writing has been a central ability for workplace success for decades (see Cooper). The U.S. Department of Labor’s *Occupational Outlook Handbook* describes the education and training required for a multitude of professions, and writing is a necessary qualification for professionals in many careers outside the humanities. For example, the *Handbook* lists the following information for engineers:

Engineers should be creative, inquisitive, analytical, and detail oriented.

They should be able to work as part of a team and to communicate well,

both orally and in writing. Communication abilities are becoming

increasingly important as engineers interact more frequently with specialists in a wide range of fields outside engineering. (“Engineering”)

The first characteristics – creative, inquisitive, analytical, and detail orientated – are traits we inculcate in FYW. Our classes specifically addresses the idea of being able “to communicate well, both orally and in writing” through semester-long discussions, reflections, brainstorming, writing, rewriting, and revising. The Department of Labor recognizes that communications skills, particularly writing, are increasingly important across the disciplines. Engineers and architects need to write proposals and submit bids, scientists write their findings in lab reports and professional journals, many specialists apply for funding in writing. Although some scholars cite the increasing importance of workplace writing as evidence *against* the FYW requirement, I believe the courses provide an essential foundation to students. Writing courses prepare students to become more aware of how language functions in their world, and provide students with opportunities to analyze and test their own theories of writing.

Methods of Placement

Placement into FYW, however, is still overwhelmingly used as a tool to filter students into appropriate courses – not as a pivotal moment in their educational experiences, or even as an integral part of the university or writing program. Basically, it facilitates the completion of a task. Its entire existence is based on assisting the user to perform a function more quickly, more easily, more inexpensively than is possible without the tool—namely putting FYW students in their proper seats. This perspective of placement has led to the proliferation of methods that sort students quickly, cheaply, and relatively accurately, with little concern by users about the messages conveyed. It has

also led to methods, often constructed within writing programs, that attempt to reclaim placement while mostly focusing on the preexisting criteria of up-front costs to universities, the time it takes to score and distribute scores, and inter-rater reliability, or how often scorers agree on a testing decision. Moreover, the specific means of placement employed by the university sends a clear and powerful message to high schools, namely teach to the test. As Hunter Breland explains,

Also related to direct assessment are issues of national impact-the message that is implicitly sent to students and teachers by direct assessment used on a wide scale: If large numbers of students are required to produce compositions for assessments important for graduation, certification, or admission to higher levels of education, then students will be encouraged to learn composition skills and teachers to teach them. (1)

Breland discusses here the difference in perception between multiple-choice tests of grammar and usage and essay tests; however, the message is the same. If high schools desire their graduating seniors to score well on the college English placement test, they almost inevitably will steer their curricula toward that test. Thus placement tests that require no writing will almost certainly influence high schools to offer English classes that emphasize multiple-choice exams and de-emphasize the difficult and often messy practice of teaching writing within purposeful rhetorical contexts.

The chart below (Figure 5) distills the most prevalent placement practices discussed in Chapters Three and Four – multiple-choice tests of grammar and usage, holistically scored essays, directed self-placement, and automated essay scoring – into those three criteria.

Placement Method	Low up-front cost?	Time?	Reliable?
Multiple-choice tests of grammar and usage	YES	YES	YES
Holistic scoring of essays	YES	YES	YES
Directed Self-Placement	NO	NO	YES
Automated Essay Scoring	YES	YES	YES

Figure 5: Traditional Criteria for Placement Practices

By the criteria above, multiple-choice tests, holistic scoring of essays, and AEE are clearly the most efficient tools to do the work of placement. Multiple-choice tests and AEE, with their essentially immediate scoring process, can even outperform holistically scored essays, which are typically scored within two minutes. AEE has the added benefit of including essays, which can increase its face validity by actually testing writing as it retains the benefits of multiple-choice tests: speed, consistency, constant availability, and low costs (Klobucar et al. 104-05). While DSP might reliably and accurately place students into their required FYW courses, it generally cannot do so quickly or inexpensively. DSP is therefore out of place in an efficiency-based system.

Unfortunately, the most efficient methods, all of which are or can be produced as both local and mass-market, large-scale exams, offer little helpful information about the ways most universities and writing programs conceive writing. The exams simply reward an easily identifiable structure in essays along with language fluency/verbal ability. If a student can produce an essay that is long enough with few language errors, it is likely she will receive a positive score on an efficiency-based test. As Ed White explains in “My Five-Paragraph-Theme Theme,” written after his own experience holistically scoring essays written for the Advanced Placement exam,

The last reason to write in this way is the most important. Once you have it down, you can use it for practically anything. Does God exist? Well you can say yes and give three reasons, or no and give three different reasons. It doesn't really matter. You're sure to get a good grade whatever you pick to put into the formula. And that's the real reason for education, to get those good grades without thinking too much and using up too much time.

(525)

In the quest for efficiency, we have removed writing from its context. In fact, Les Perelman has termed such tests to be “bullshit” because students are rewarded for including irrelevant or inaccurate information to answer short essay questions that have nothing to do with their knowledge or experience domains. This, of course, is in stark contrast to most testing or writing situations, where we are expected to demonstrate knowledge based on course texts, assignments, and discussions or professional expertise. Such placement tests work to simply shuttle students to writing classes, fulfill their writing requirements, and be “finished” with writing.

AEE technology currently identifies features of writing that include:

(a) grammatical and word usage errors (based on native and non-native speaker writing data), (b) mechanics errors (e.g., *spelling*), (c) presence of essay-based discourse elements (e.g., a thesis statement), (d) development of essay-based discourse elements, (e) style weaknesses (e.g., *overly repetitious words*), (f) a content vector analysis feature comparing an essay to the set of training essays, (g) a content vector analysis feature comparing an essay to the set of training essays that received a score of 6,

(h) average word length, and (i) a word frequency-based feature. (Burstein 206)

It is significant to note that there is no mention of the development of ideas or accuracy of information. While it would be difficult for a machine to analyze and score how a piece of writing develops – not merely how it supports a thesis statement – it should be relatively easy for a machine to fact-check. The software, however, is not designed to do so; consequently, students can submit and receive high scores for writing nonsense. In an article discussing how he has been able to trick AEE tests, Perelman uses an extreme example:

For a question asking students to discuss why college costs are so high, Mr. Perelman wrote that the No. 1 reason is excessive pay for greedy teaching assistants.

“The average teaching assistant makes six times as much money as college presidents,” he wrote. “In addition, they often receive a plethora of extra benefits such as private jets, vacations in the south seas, starring roles in motion pictures.” (Winerip)

Perelman’s essay was rewarded, even as it included this blatantly false statement. Similarly, students are not penalized for incorrect facts because it is a test of their writing, not of their knowledge (Perelman “Mass-Market”). Unfortunately, in most situations, writers need to create texts that do make sense and contain some semblance of truth.

The problem with limiting test criteria to “quick, cheap, and reliable” is that once a test-taker encounters writing in a FYW curriculum, or in most places across the university, or even in their professional or personal spaces, the test is perceived as being

irrelevant. Such tests “require capabilities that are irrelevant to the intended interpretation” and “result in performances that under-represent or provide a narrower interpretation than what is intended” (Messick qtd. in Moss “Testing” 114). The tests provide extremely limited information about students’ writing ability, so limited, in fact, that it is difficult to extrapolate how well a student is prepared for a rhetoric-based writing curriculum. This is not surprising when the test that places the student limits her to writing a response that contains about 80 words written in thirty minutes – which might otherwise be deemed a paragraph (Perelman “Critique”). Many teachers and parents can attest to the typing abilities of their college-aged students on handheld electronic devices as well as computers. In a testing situation, however, their quick typing cannot make up for reading a topic they have had no previous experience considering or working with, developing a position on it, brainstorming, drafting, writing, revising, and editing in thirty minutes.

Language Difference in the University

I’d like to return to the idea of context and how it relates to our increasingly multilingual students. Terms such as *english*, *englishes*, and *World Englishes* are frequently used in scholarship and in society to mark the evolution of English from being a monolith to becoming a world language. English is not only the dominant language in countries like Great Britain, the United States of America, Canada, or Australia; it is also an official language in countries including India, Nigeria, Pakistan, and Trinidad, and is the world language for business. Campuses across the U.S. are increasing their international recruitment efforts, and immigration advocates are liberalizing admissions policies to provide higher education opportunities for children of undocumented

immigrants. In the midst of this era of difference within one language (English), and increased linguistic diversity nationwide, universities and colleges still overwhelmingly use assessment technologies that are built on the notion of a unified English, a “Standard Written English” or “Edited American English.” In many testing situations – multiple-choice tests of grammar and usage, or essays scored by AEE – it is presumed students are familiar with the conventions of EAE. They are penalized for not proving that knowledge. This view of language as static, as something that ought to be used in one way, is shared by those who propose legislation for English as a national language. A shocking similarity exists between legislators who recall the United States’ diversity but propose laws that penalize multilingualism, and universities that include diversity missions in their strategic visions but implement methods that punish students for being diverse and for not adhering to an outdated and harmful view of language use.

As linguists and writing teacher-scholars have made clear, however, language is dynamic, and

notions of the “Standard English Speaker” and “Standard Written English” are bankrupt concepts. All speakers of English speak many variations of English, every one of them accented, and all of them subject to change as they intermingle with other varieties of English and other languages. Likewise, Standards of written English are neither uniform nor fixed. What constitutes expected norms – for example, Edited American English – varies over time and from genre to genre. Indeed, these genres themselves change boundaries and intermingle. (Horner et al. 305)

There is no one form of English or even of written English: the standards change dependent on the writing situation and the era. Professional writers and teacher-scholars make use of new standards for writing in contexts like live-tweeting a conference. In doing so, writers have only 140 characters to record their impressions, so they must be creative in their writing, punctuation, and abbreviation. Time is of the essence; one must type quickly to capture a snapshot of a presentation and pay attention to the information shared next. Spelling accuracy is less important than the message. Moreover, messages that are retweeted and recirculated are valued, irrespective of the errors they exhibit. Context is determinative of content and even form, and over time as well language changes, just as do standards for writing. A modern example we have seen become transformed in the past few years is usage of the plural pronoun “their.” Consider the statement, “The writer showed their bias in the article”. Language purists argue the pronoun “her” or “his” would be more accurate. Everyday language users, however, are no longer content using the male pronoun in situations where a person’s gender is unknown, and they find the phrase “her or his” to be wordy, and “s/he” awkward in their writing, let alone their speaking. Many have taken to using “their” as a mark of gender neutrality in their everyday speech and writing, and it is becoming more widely accepted in academic and professional settings as well, despite the fact that such usage violates prescriptive rules of number agreement.

Composition scholarship consistently argues for the inclusion of diverse perspectives and values – beginning with pedagogies inspired by the open-admissions policies of the 1970s, the 1974 Statement to affirm “Students’ Rights to Their Own Language” (which was reaffirmed in 2003 and 2006), the social and linguistic turns in

writing studies, and the current emphasis on translingual writing (Lunsford and Ouzgane 2). SRTOL states:

We affirm the students' right to their own patterns and varieties of language – the dialects of their nurture or whatever dialects in which they find their own identity and style. Language scholars long ago denied that the myth of a standard American dialect has any validity. The claim that any one dialect is unacceptable amounts to an attempt of one social group to exert its dominance over another. Such a claim leads to false advice for speakers and writers, and immoral advice for humans. A nation proud of its diverse heritage and its cultural and racial variety will preserve its heritage of dialects. We affirm strongly that teachers must have the experiences and training that will enable them to respect diversity and uphold the right of students to their own language. (2-3)

CCCC and NCTE value the diversity of their students, celebrate the range of experiences that inform our country and our classrooms, and seek to protect linguistic difference.

SRTOL highlights the indisputable relationship between language and identity, a concept regularly explored in some of the most lasting scholarship in Rhetoric, Composition, and Writing Studies. In *Voices of the Self*, for example, Keith Gilyard explains how identity and language are inextricable. He details the “clash of value systems” he encountered as a young man who was essentially forced to deny his home dialect and identity to succeed academically in a predominantly white setting. Gilyard argues: “A pedagogy is successful only if it makes knowledge or skill achievable while at the same time allowing students to maintain their own sense of identity,” and he questions why students should ever be

asked to forsake their identity for the sake of learning (11). In the context of placement procedures used to determine students' writing abilities, we should extend Gilyard's discussion to question why students are penalized for using their home dialects, why our requirements seek to suppress their identities or mark them as deficient.

bell hooks discusses similar ideas in her *Teaching Community*, a text based on a pedagogy of affirmation and respect (103). hooks explains that for students to function well in our nation built on democracy and diversity, they need to be able to function in a range of settings (80). Bruce Horner, Min-Zhan Lu, Jacqueline Jones Royster, and John Trimbur expand on the idea: "even if we were to accept that our students are English monolinguals, they are unlikely to be restricted in their writing, or their speech, to audiences of only other English monolinguals" (311). Interestingly, the U.S. Department of Labor states a similar belief in their *Occupational Outlook Handbook*. Most professionals need to demonstrate oral and written communication skills as they are required to interact across disciplines, communities, languages, and countries. Composition scholars and federal economists agree that to instill successful habits of mind in students, to help students develop into successful participants in their communities, teachers should expose students to a wide range of ideas and beliefs (see Appendix G). Universities value this responsibility of developing cultural awareness across campus, evident not only in local, regional, and national guidelines for student learning outcomes but also in their mission and vision statements. The chart below documents just a few examples of universities nationwide that articulate their dedication to diversity.

Location	Mission/Vision Statement
Urban, Midwestern, public	“Diversity in all of its definitions, including who we are, how we think, and what we do.
Urban, Midwestern, private	“_____ is the school of choice for those who wish to seek new knowledge in the service of humanity in a world-renowned urban center as members of a diverse learning community that values freedom of inquiry, the pursuit of truth and care for others.”
Urban, Midatlantic, public	“_____ mission is to provide a diverse population with an excellent university education.
Urban, Western, public	“Embrace a culture of diversity, internationalization, and inclusion;”
Suburban, Western, private	“We cherish our diverse and inclusive community of students, faculty, staff, administrators, and alumni, a community that is enriched by people of different backgrounds, respectful of the dignity of all its members, enlivened by open communication, and caring and just toward others.”
Rural, Southern, public	“_____ provides an academic experience that emphasizes critical thinking; encourages intellectual depth and creativity; challenges and inspires a diverse community of undergraduate, graduate, and professional students; provides enriching opportunities outside the classroom; supports lifelong learning; and develops a sense of global responsibility.”
Rural, Northern, private	“_____ embraces diversity with the knowledge that it significantly enhances the quality of a _____ education. _____ recruits and admits outstanding students from all backgrounds, regardless of their financial means.”

Figure 6: University Mission and Vision Statements

University mission and vision statements are public professions of the core principles that guide the institution. Colleges and universities welcome diverse populations and articulate a sense of responsibility in interactions among local and global communities. It thus follows that composition instructors should also welcome and honor students' linguistic, cultural, gender, sexual, religious, ethnic, and racial diversity. That is, our work should demonstrate a dedication to the democratic ideals of the United States, especially the concept of a national identity stemming from the diversity of cultural heritages evident in the national motto, "Out of many, one." The dedication to diversity expressed in universities nationwide may view the motto as out of many people, one nation or one university: we all work together to create a strong space that respects, values, and draws on our diversity. Employing placement tests that deny, suppress, or penalize students' diversity effectively dissolves our culturally rich heritage for the sake of one national identity, vision, and language – typically represented as a white, middle-class, monolingual English-speaking American identity. Tests created in the service of efficiency do not allow for language difference, language changes, or language flexibility.

Placement in FYW

In 2014, the efficiency model of placement is no longer adequate for use in universities. Placement clearly does more than determine which quantifiable writing traits a student can demonstrate. Therefore, a method of placement must go beyond segregating students based on a limited understanding of writing ability. In addition to predicting which FYW writing course a student is best prepared for, the results of a placement decision imposes various institutional and social labels on students – remedial,

mainstream, advanced, etc. – most of which will identify them throughout their academic career. Students at either extremity of the placement spectrum are tracked by various bodies across the university, including the writing program, writing center, demographics offices, and honors colleges. In many instances, because of the results of their placement procedure, students are *locked into* a path until they graduate or leave the university. Furthermore, as explained in more detail in Chapters One and Two, placement communicates essential information about the university to its wide range of stakeholders. As a result, writing assessment practitioners (Huot; O’Neill, Huot, and Moore; White; Williamson; Yancey) have called for methods of assessment and of placement that are site-based, locally-controlled, context-sensitive, rhetorically-based, and accessible (see Appendix A). Huot “propose[s] that we design the complete assessment procedure upon the purpose and context of the specific writing ability to be described and evaluated” to “take charge of how are students are to be evaluated....to build and maintain writing assessment theories and practices which are consonant with our teaching and research” ((Re)*Articulating* 102; 108). Just as curricula are typically developed within individual writing programs, the argument goes, so should the methods used to assess student work within those programs.

O’Neill, Moore, and Huot discuss the vital importance of context and rhetorical sensitivity in creating good assessments. The assessment practitioners argue:

Because writing assessment is fundamentally about supporting current theories of language and learning and improving literacy and instruction, it should involve the same kinds of thinking we use every day as scholars and teachers. (59)

Their argument relies on the idea that principles of good writing pedagogy can guide good writing assessment, that all forms of assessment – including placement – can uphold the field’s deepest values about writing. O’Neill, Moore, and Huot are even more explicit about how curriculum and instructional choices are led by our writing values and are dependent on context:

In much the same way, context informs the decisions we make as teachers. We consider not only what teaching methods are available but how they coincide with the mission of the school, whether they support the goals of a particular program or course, and how we will modify them for different groups of students. (59)

Writing program administrators, writing teacher-scholars, and writing assessment practitioners value writing that reflects students’ rhetorical awareness and knowledge. Attention to context is a central aspect of that awareness, and to successful writing in any genre. Figure 7 allows us to reconsider the most prominent methods of placement – multiple-choice tests of grammar and usage, holistically scored essays, DSP, and AEE – in light of third-wave assessment criteria. Specifically, using the concepts of validity, context-sensitivity, and the representation of core values provides us with a much different perspective than using the traditional criteria in Figure 5.

Placement Method	Valid?	Context-sensitive?	Core Values represented?
Multiple-choice tests of grammar and usage	NO	NO	NO
Holistic scoring of essays	NO	NO	NO
Directed Self-Placement	YES	YES	YES
Automated Essay Scoring	NO	NO	NO

Figure 7: Modern Criteria for Placement Practices

As is evident, there is no overlap between the traditional criteria used to evaluate placement practices and modern criteria. Consequently, Directed Self-Placement becomes the most suitable method to place students into FYW, whereas it is the least suitable method with regards to time, cost, and efficiency. It is an attractive alternative to mass-market tests because its success is totally dependent on its relationship to a local writing program. Directed Self-Placement is clearly one method that is organically developed out of the needs of a specific site of writing in a specific university (Broad). As Gere et al. demonstrate, establishing validity within DSP systems might not be easy, but it is certainly possible. Compositionists identify DSP as an impressive alternative to the mass-market dominated placement market because its purpose is to adhere to local curricula and to provide opportunities for students to assert their agency; that is, to demonstrate some of the most central values about writing pedagogy.

Betty Bamberg crystallizes the central goal of writing assessment: to devise methods that allow students to demonstrate their competence as writers, *based on the course content*. Here is where validity is essential. If, in an American History course from 1792-1860, the final exam included questions about the 1898 Spanish-American War,

there would be an outcry! Instructors are expected to examine students based on what they teach. The articulation of learning outcomes binds us to teach those outcomes, and to assess student work based on those outcomes. Similarly, I believe that we are bound to place students based on our understandings of the outcomes and of the curriculum a student is preparing to enter. If, as a field and as a community of scholar-teachers, we have agreed to assess students based on course content, why wouldn't we also place students based on their preparation to begin working on that content? If Composition Studies has disproved the usefulness of using multiple choice tests of grammar and usage to test writing ability since they disproportionately penalize students who speak linguistic minority dialects as well as students who are labeled exceptional in high school and haven't had a grammar class since the seventh grade then why would our writing programs not only consent to using them to place students into their FYC courses but *advocate* for their continued usage? The only way a multiple-choice test of grammar and usage can be valid is if "evidence and theory support the interpretations of test scores" (*Standards*). Simply put, rhetoric, composition, and writing studies scholarship *does not* support the interpretation of such test scores to represent students' competence as writers who can respond to their rhetorical situations. To create a valid placement procedure, administrators must carefully consider how they will interpret and use the results, and whether those interpretations and uses reflect their curriculum and student learning outcomes.

Paying careful attention to a placement procedure's validity and designing it to be context-sensitive, however, does not necessarily guarantee fair, ethical, or even legal treatment of local populations of students, as a recent study by Mya Poe, John Aloysius

Cogan, Jr., Norbert Elliot, and Tito G. A. Nurudeen, Jr. makes clear (Poe and Elliot). At “Brick City College,” African American students were disproportionately placed into basic writing courses, based on a locally-designed placement test and *in spite of* their SAT Writing test scores that were higher than state and national averages. Their mean SAT Writing scores were also, on average, higher than Hispanic students’ scores; however, 47% of African American students were required to take basic writing, compared to 28% of Hispanic and just 10% of white students. Poe et al.’s study on disparate impact discrimination demonstrates the necessity of examining our writing assessments in light of our student populations: “Whether the test is developed locally or commercially, disaggregated information on student performance, understood in terms of gender and race, must be provided before the test can be meaningfully used” (Elliot et al. 31).

Neo-empiricism and Dialogue

Paul Walker highlights some of the problems associated with methods of placement that develop out of writing studies as opposed to educational measurement:

We are implicated in current trends because we have publicly affirmed “best practices” in assessment that, even when insisting that context matters, are method-heavy, meaning they are easily misused to contradict that guideline. For institutions under pressure to “show results” immediately, the quickest and most inexpensive move is to ignore context and adopt a “proven” method. (Walker para. 4)

Although practices like DSP are valid, context-sensitive, and uphold the field’s values about writing, they are perceived as being insufficient in an era more concerned about

financial bottom lines than careful and ethical practice. Returning to Figures 5 and 7 and their differing criteria for evaluating placement procedures, it seems as though the century-old tension between educational measurement and academic communities is irresolvable. Some scholars, including Pamela Moss, Ed White, Norbert Elliot, and Carl Whithaus, believe assessment specialists from academia and from educational measurement can “enlarge the dialogue,” collaborate, and find solutions to the placement problem. Others, like Paul Walker, Anthony Petruzzi, and Chris Gallagher explain how in this era of “neo-empiricism,” psychometricians hold so much power that a dialectic is essentially impossible. Turning to Sharon Crowley’s discussion of rhetoric and fundamentalism, however, provides insight on how to drive through this impasse to create better assessments. Crowley suggests revising methods of argumentation when encountering positions based on “foundationalism,” or an ideology “taken to be noncontingent – that is...taken to apply noncontextually or universally” (*Toward* 13). She explains,

To attack a fundamentalism on the ground that it is not rational is to apply a standard that is valued in some belief systems and not in others; to treat rationality as universally binding is, willy-nilly, to fall into yet another fundamentalism. (*Toward* 13)

To Crowley, Christian fundamentalists and liberals cannot enter into civil discourse because their foundational ideologies are so disparate *and* because they use their individual value systems to view the other system. It is therefore ineffective to use traditionally held liberal beliefs to argue against fundamentalism because fundamentalism does not value those same beliefs. The two groups are simply talking to themselves: there

is no crossover. Academics are often charged with the same fault; even when we blog, for instance, we are only reaching other academics, not our publics (Thomson and Mewburn). Crowley's solution to work toward a civil discourse is to uncover traditional methods of rhetoric to appeal to different audiences: "Since antiquity rhetorical theorists have understood the centrality of desires and values to the maintenance of beliefs" (*Toward* 4). Rhetorical argumentation can work, Crowley claims, because it provides us with a variety of strategies and appeals to use when we encounter ideological difference.

Writing teacher-scholars, WPAs, and even assessment practitioners have been unsuccessful in creating a dialectic with psychometricians, university administrators, and our publics because we represent two totally different ideologic positions. While the first group of academics are foundationally concerned with improving teaching and learning, the second group of professionals and publics are most focused on efficiency. Teachers resist multiple-choice tests, holistically scored essays, and AEE because they fulfill the standards of time, cost and reliability and embrace DSP because it represents validity, context-sensitivity, and core writing values. The two groups can keep discussing methods of determining writing ability for another hundred years, but it does not appear that their positions will ever converge.

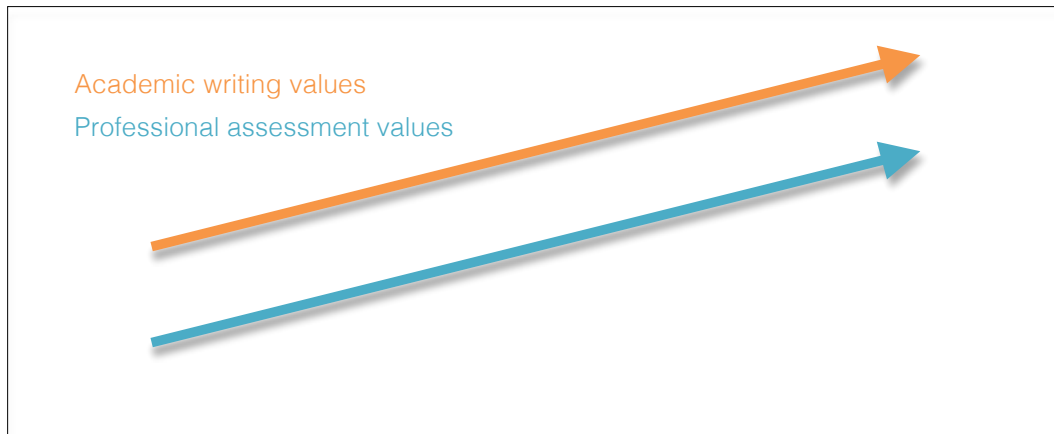


Figure 8: Divergent Ideologies

I believe a dialogic perspective of placement, however, can provide some suggestions about how to bridge the seemingly constant bifurcation of the two assessment communities.

A Dialogic Perspective of Placement

A dialectical approach to placement or to assessment has been insufficient. As Gallagher explains,

[Dialectical approaches] provide no means by which to assign priority or preference to any set of perspectives and positions over any other set. A rhetoric operating in this way within an inequitable scene only reinforces the current order, especially when that order—shaped by the perspectives and positions of those in power—is understood as “common sense.”

(“Being There” 461).

A dialectic approach is ineffective within placement because one group – psychometricians, university administrators, and the concerned public – is situated within

a hegemonic discourse. From their position of power, the language of educational measurement and of neo-empiricism appears to be natural within public discourse (Crowley *Toward* 12). There can be no dialogue without a semblance of equal or shared power, however, and as Gallagher describes, writing assessment is an enterprise with a distinctly hierarchical construction, as evident in the graph below (Figure 9).

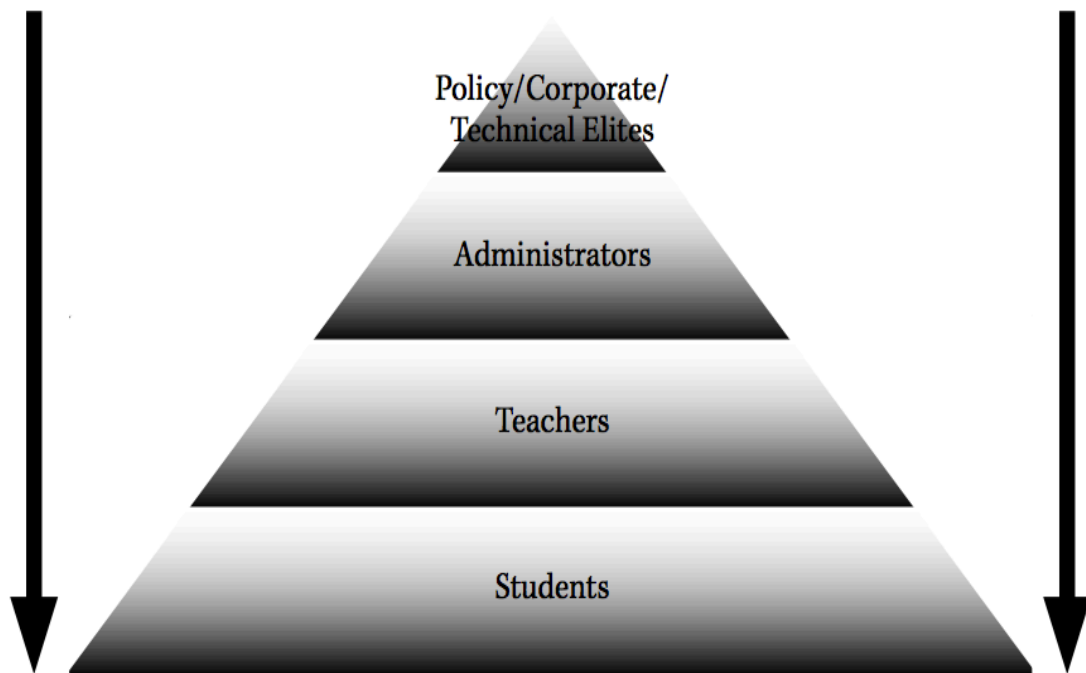


Figure 9: Chris Gallagher's Bureaucratic-Institutional Model of Authority

Despite the numerous calls from within academic spaces to reach out and collaborate with those in power, the lasting effect on writing assessment and placement has been minimal because of the bureaucratic-institutional model of authority. In our classrooms, for instance, we may listen to student pleas for fewer assignments, but that does not mean we will fulfill their wishes. When instructors or WPAs decrease the number of writing assignments, for example, it is after much deliberation and only occurs when the calendar

will allow for it and the omission will still prepare students to fulfill the articulated learning outcomes for the course. That is, we listen to students when their concerns align with our standards and values. The same relationship exists as we move up the hierarchy. Simply put, since those who are situated in the upper echelons of the model hold more power than those below them and get to make the decisions. Dialectic, in this case, does not work because of the unequal power differentials. Alternately, a dialogic perspective, which is reminiscent of Crowley's alternative to the fundamentalist/liberal standoff, can provide realistic guidelines to work toward better placement practices.

Language and Power: Validity

To achieve a dialogic among the various stakeholders involved in placement, we must first work toward a shared use and understanding of the language used. That is, I believe writing teacher-scholars, WPAs, and writing assessment practitioners should continue to use the language of educational measurement that has been in place for a century. This is not a radical position, and to some, like Patricia Lynne, it is seen as welcoming our own subjugation. While several terms have been offered as alternatives to “validity” and “reliability” – including meaningfulness, fairness, ethicalness – I maintain “validity” and “reliability” are still the best terms to use. “Validity,” for instance, has cultural, interdisciplinary, and institutional cache. Patricia Lynne explains her resistance to using these vestiges of positivism:

When prominent assessment scholars such as Brian Huot and Edward M. White argue that psychometric principles are appropriate for writing assessment and that compositionists need to understand the terms of educational measurement theory in order to justify their practices to those

outside of composition, I feel they put our practices and our principles at risk. (7)

I am arguing just the opposite: without using the terms accepted across disciplinary boundaries, we put our practices and our principles at risk. It is only by using a shared language that those of us concerned about the teaching and assessing of student writing can do the work our expertise suggests is best for our students. If the field used terms other than reliability and validity, we would lose the essential ability to communicate across social boundaries. “Validity” is something anyone with a cursory understanding of statistics is familiar with, including many of our writing students, since statistics is often a high school or college mathematics requirement. Even the outdated definition of validity – something measures what it says it will measure – implies people will have to interpret results to determine if the test is being used appropriately. It is not difficult to extend the definition or explain its current understanding to include the idea of accurately or appropriately using test results.

Cindy Moore insightfully explains the precarious position of a WPA and of writing faculty, and how using ambiguous, field-centric terms may, in fact, reduce our efficacy:

From a feminist standpoint, my identification with discipline or department may be destabilized further if I am the only woman in the room and the only faculty member who is a member of a traditionally feminine or feminized discipline (like Composition). As several female writing program administrators have convincingly argued, once I step my female body out of my program, out of my department, and into another

context, with its own set of histories and expectations, I cannot assume that my expertise or my language will be interpreted in the same way (see, e.g., Barr-Ebest, 1995; Miller, 1996, Schell, 1998). In fact, from a sociolinguistic perspective, any attempts to speak differently without careful consideration of the particular context might actually work to concretize perceived marginal status. (470)

While scholars like Patricia Lynne argue against using the term “validity” because of its association with the positivist tradition, it is precisely *because* of this tradition that it holds such weight in our cultural, interdisciplinary, and institutional conversations. If WPAs were to use a different term, like Lynne’s “meaningfulness,” we would lose credibility with the very people with whom we need to establish it.

To Moore, language can represent and reinforce social power structures, a concept reminiscent of James Berlin’s writing that has become central to the field of Composition:

Language in all its uses structures rather than simply records experiences. Thus, language never acts as a simple referent to an eternal, extralinguistically verifiable thing in itself. It instead serves as a terministic screen to form and shape experience. ... In other words, language is a product of social relations. (Berlin 34)

Berlin’s work reminds us that we are all co-creators of knowledge, who participate in and shape our communities – including our university and social communities – through writing and language choices. Choosing not to use a dominant term in public conversations with university administrators, who may hail from architecture, for

example, further marginalizes the work of compositionists and writing program administrators. We may be perceived as too disconnected from “the real world,” too flighty to understand foundational concepts of statistics, too feminine or feminized. Ultimately, Moore explains, choosing not to use the accepted terms of reliability and validity can cause WPAs to lose status – which can lead to decreased control of our programs and diminished views of our expertise. I believe it can also increase ideologic impasses and decrease communication across all those who are involved in placement. Choosing validity to be the central term to discuss placement in universities indicates our dedication to create and uphold methods that are coherently aligned with our writing programs, that communicate our best values about writing, that treat students with respect.

Language and Power: Numbers

In addition to resisting psychometric terms of assessment, those in academic writing communities also often resist the numbers used in testing. Haswell summarizes the “fear of numbering” as such:

Fear of numbering is widespread and chronic. In composition studies, for instance, the case against number-based scholarship has been made and remade. Only trivial things can be measured. The relationship between rhetorical values is not mathematical. Data analysis breaks language effect into pieces and does not reassemble it into a whole. Numbers themselves are abstract, not real, and research findings in the form of unbodied averages and correlations cannot be transferred to the living bodies of students in classrooms. In testing, numbers only simplify complex abilities

and performances, falsely rationalize the intuitive, and encourage comparison of academic and demographic contexts that are singular, not comparable. (“Number” 413)

While there are many reasons teacher-scholars have traditionally resisted numbers about writing, the most prominent is that quantitative data provides a limited perspective of students’ writing abilities. For many, numbers can only provide an abstraction of the complexities of writing. Haswell argues, however, that the more WPAs use numbers and data within their programs, the more able they will be to “stave off outside assessment” (414). Numbers can be powerfully convincing, especially in a discourse community that values objectivism and neo-empiricism. Haswell claims we should “fight numbers with numbers” and be prepared with quantitative data to be more persuasive (414). I agree with Haswell, and with White that the more we can do assessment, the more we can do with assessment.

By way of caution, however, I turn to Bill Condon’s discussion of AEE to offer a more nuanced understanding of how to use numbers:

the score[s] on a large, commercial assessment instrument...are subject to the fallacy of surrogation – the substitution of a statistical artifact – a number – in place of the need for complex information. For the most part, these assessments...yield a number that, quite simply, does not reveal much information beyond the number and generally, a percentile ranking of one test-taker among all the test takers. (“Red Herrings” 101)

A score tells us nothing about a student. To become meaningful, that score must be associated with a percentile to compare how students do in relation to each other, or to

predict how students might fare in a course, in which case the testing center or administration must determine what score ranges relate to which courses. To Condon, the scoring that happens on mass-market tests is context-less and must be interpreted and reinterpreted by local administrators and teachers. At UW—Milwaukee, for example, students take a University of Wisconsin System standardized multiple-choice test of grammar and usage. The UW Testing Center determines the weight of each question on the test, and students earn a “raw score,” which is a result of the points they earned or lost on each question. The raw score is then translated into a scale of 0-4 as determined by each local institution, which then is correlated to a FYW course or exemption from FYW altogether. The local institution determines the cut scores, with advice from the university system testing center, as well as the placement codes that relate to a FYW course (see Figure 10).

Placement Test Raw Score	Placement Code	First Year Writing Course
150-314	0	English 095
315-384	1	English 101
385-464	2	English 101
465-524	3	English 102
525-850	4	Requirement satisfied

Figure 10: UW—Milwaukee English Placement Test Score Chart¹⁴

¹⁴ Adapted from “English Placement Test and Composition Courses,” prepared April 2012 by UWM FYW Program

Before 2005, for instance, students who received a placement code of 0 or 1 were channeled into developmental writing courses. After those courses were revised, more students were mainstreamed into English 101, which is now required for students who earn a 1 or 2 placement code. Similarly, because fewer students had been enrolling in developmental courses, local WPAs shifted the raw score cutoff to a higher number to target more students in need of supplementary instruction (Brusin). Large-scale tests contain multiple instances of translating numbers, including determining points on a test, converting that to a “raw score,” translating that into a “placement code,” associating placement codes with FYW courses, and determining how raw scores relate to placement codes. This serves as an example of one way writing teacher-scholars are already immersed in numbers and are using them, although sometimes inefficiently.

In addition to quantitative data, the most important data for those interested in placement are numbers concerning cost. Mass-market tests have proven to provide low up-front costs to universities. Administrators and publics believe that the more a placement method has to do with writing, the longer it will take to score, and the more money they will have to spend. Condon and others highlight the importance of *cost-effectiveness* rather than simply cost-efficiency (“Future”). A cost-effective placement method might be more expensive up-front, but it would save writing programs and institutions money in the long run. One way to see the cost savings is by using methods that result in fewer students challenging their placements or requesting to switch courses; another is by shifting the decision-making from the university or writing program and onto the students, as in DSP. If the focus on placement methods shifts from being quick to being accurate or valid, for instance, it can help to decrease funds otherwise used when

students are placed in courses for which they are over- or under-prepared. The iMOAT online challenge process at Louisiana State University is one example of a cost-effective placement method. LSU uses student scores from the ACT or SAT to determine an initial placement (which means students cover the cost of placement). Those who disagree with their FYW course placement are able to challenge their scores and complete a rigorous DSP profile, which includes a typical assignment from the writing course. Only about 4-5% of an incoming class submits an essay for the online challenge, which means the cost is still quite low for WPAs and instructors to review student disputes in this hybrid model. Creating a cost-effective placement method means the onus is on WPAs, writing teacher-scholars, and writing assessment practitioners to use their rhetorical savvy and quantitative data to persuade their otherwise resistant stakeholders. By attempting to relate to those who control power or money in our local and national contexts, it might be possible to create placement practices that we believe are ethical and sound, that can appeal to publics concerned about empiricism and cost, and that can thrive under difficult constraints.

Conclusions and Implications

In an era of state- and federally-imposed furloughs on public employees, slashed budgets, increasing class sizes, and shrinking instructor pools, writing program administrators are already overburdened as they fulfill their programmatic and institutional duties. It is my contention, however, that we need to include placement practices in those regular duties, to confront the institutional lore and thoughtfully investigate whether our practices really are “good enough” by 1) adapting methods and premises from multiple disciplines and stakeholders, 2) collaborating with multiple

disciplines and organizations across our institutions, and 3) understanding the deep and multiple meanings within hegemonic discourse to more effectively persuade our skeptics.

Huot writes that we ought to be “attempting to assess a writer’s ability to communicate within a particular context and to a specific audience who needs to read this writing as part of a clearly defined communicative event” (“Toward” 559). To me, this is an explanation of why assessment should be rhetorically-based: students should be aware of the exigency for their writing, and how their writing is both affected by and affects their audience. I believe placement is indeed a rhetorical act that occupies a central space within a university and that communicates a multitude of messages to a range of audiences. Consequently, placement deserves to be carefully considered by writing program administrators, writing teacher-scholars, and writing assessment practitioners in addition to university administrators, students, and members of our concerned publics. Placement has a number of purposes beyond attempting to identify students’ writing abilities: it can be used to advance the teaching and learning that takes place within writing programs; it can prepare students for their upcoming university experiences with writing; it can improve communication between academia and the educational measurement community, our various publics (i.e. parents, students, corporations, governments, media), and our colleagues in K-12 education.

Placement procedures historically involve multiple units on campus, not only one writing program, and I believe tapping into these relationships can help strengthen our methods of assessment. There are testing centers, of course, in addition to offices of institutional effectiveness that track assessment efforts across campus; student success centers dedicated to offering resources particularly to students in remedial courses;

budgetary and planning departments; academic advising; and offices of institutional research, focused on recruitment and retention. David Blakesley found that advisors were happy to implement DSP after he met with them because they had previously never been included in campus-wide decisions; he recognized their role within the placement process, and particularly within the ability for DSP to be successful on their campus. In my own work at UW—Milwaukee as Coordinator of Basic Writing I found a similar lack of communication between the writing program and locations like the Accessibility Resource Center, Academic Opportunity Center, and Student Support Services. It is difficult to build bridges between various departments, programs, and colleges, as well as between academic and support services in a large state university – but it is not impossible. I believe we can create better placement methods by recognizing the range of participants involved in placement, and by working with them to improve the teaching and learning conditions for students. I therefore am not calling on teacher-scholars, writing program administrators, and writing assessment practitioners to reframe the conversations about placement; I am not convinced it is possible to do so across universities and within public discourse. Time and cost, in addition to reliability and validity, are still prominent concerns for many participants. If we are to re-envision placement in the twenty-first century, we need to be able to effectively communicate with each group affected by the process, even those who are in power.

So what can be done from a seemingly powerless position as an instructor, a teacher-scholar, even a graduate student? Of course, one way to alleviate the time, cost, and ethical burdens of placement is to require a two-semester First-Year Writing sequence for all new students, and to limit enrollment in each section to 15 students. This

is probably not a viable option in most locations, however, so instead I offer two suggestions, using my own campus as an example.

Challenges to Placement Decisions

Within the University of Wisconsin System, each of the twenty-six campuses decides on their placement practices to ensure relevance on each campus, where populations of students and curricula differ. While some locations use ACT scores or writing samples to place students into their required courses, most use the Wisconsin English Placement Test (EPT). Those that use the EPT also receive research support from the UW Testing Center, and are encouraged to incorporate multiple methods of placement to ensure accuracy. The Center works with each campus to determine “cut scores” that place students into various course levels, based on conversations about local curricula and FYW programs. The University of Wisconsin—Milwaukee currently uses the EPT because it is efficient, but students and teachers are surprised, confused, and even angered (as in my Fall 2012 developmental writing course) each semester by placement decisions across the FYW sequence. One way to mitigate some of the confusion while also adhering to best practices within writing pedagogy and assessment is to allow for challenges to placement decisions, initiated by the student or by the WPA.

The process at Louisiana State University serves as a good model for student-initiated challenges. At UW—Milwaukee, students could continue to take the EPT and decide whether to dispute their score once they receive it at their summer orientation session. Alternatively, instructors could confer with students about the possibility to challenge their placement decision after the first class session wherein students are required to submit a brief piece of writing in response to course materials. While we do

not currently have an online system in place, it is possible one could be developed through e-Portfolio, which the university uses in its online content management system, and which would help establish security in student submissions.

The key in LSU's model is for students to have an extended period of time (at least one week) to review several articles and write a response based more on inquiry and exploration than on a one-sided argument. Students are evaluated on their ability to report on and interpret texts similar to those they can expect throughout the FYW sequence. A UW—Milwaukee challenge could use an assignment aligned with English 101, the course most students place into. WPAs could create new assignments or could even draw on curricula from the recent past since the FYW program keeps electronic and print records of all instructors' curricula, most course texts, and many individual assignments.

Ed Jones' conclusion that most students who score in the middle on placement tests, even on large-scale standardized tests, will be adequately served in their first FYW course inspires my next suggestion. He claims students who score on the peripheries, who are placed into developmental writing or who are exempted from their writing requirements, are those whose scores may need a closer look ("ACCUPLACER" 113). I suggest WPAs invite students who might otherwise be seen as "outliers" to challenge their placement decision by the process outlined above. Alternatively, WPAs could work with admissions personnel or data to verify the decision. Many students nationwide, particularly at institutions that employ a holistic admissions system, are required to submit high school transcripts, standardized test scores, letters of reference, and a personal statement, which are uploaded into admissions databases. WPAs do or could have access to the databases, and could work with their administrative team or a FYW

committee to make decisions about students who would place into development writing or out of the writing requirement altogether. Collaborating on decisions would lead to another professional development opportunity for faculty, and another chance to discuss the writing program's curriculum and theoretical foundation. It would also align with educational research that suggests students' past experiences are good (and in some instances, better) indicators of their future university performance than their standardized test scores might indicate.

Next Steps

This project has set the stage for the next logical step in the research process: using multiple methods, including “surveys, analyses of students’ essays, class observations and interviews, experimental, or quasi-experimental studies” to analyze the procedure I have laid out (Bamberg 29). Students are at the center of our careers as instructors and administrators. I believe it is essential to work with them and to understand how our various procedures affect them, rather than making assumptions (even well-informed assumptions) that often end up being counter-productive for our students, ourselves, and our institutions.

My focus here has been on how students are placed into mainstream First-Year Writing classes, but there is certainly room to examine how multilingual students operate within FYW programs, and to examine how viable courses like remedial or basic writing are, particularly as open-access admissions policies increase in urban areas throughout the country. At the University of Wisconsin—Milwaukee, for example, we have significant populations of Hmong and Hispanic students, in addition to students who speak Arabic and a variety of Eastern European languages as well as English. Many of

these multilingual students place into developmental writing. Analyzing their experiences with placement, both into composition classes within FYW and within the English as a Second Language program, would be valuable. An examination into how admissions policies have historically influenced placement practices and decisions would also be worthwhile; as Kathleen Blake Yancey's foundational article reminds us, it is essential to "look back as we look forward."

I am convinced we need to continue creating and supporting practices that fulfill our ideals about writing, and that communicate our core values about the teaching and learning of writing. When it is not in our power to create new methods, I believe we can resist invalid methods by using quantitative data and terminology others – including university administrators – can relate to. I believe we can come together in a position of mutual respect and shared understanding from educational measurement and academia to work toward practical alternatives. Ultimately, I believe that working toward understanding the various values within placement can only lead to stronger methods. I am eager to continue that work.

Works Cited

- Adler-Kassner, Linda, and Peggy O'Neill. *Reframing Writing Assessment to Improve Teaching and Learning*. Logan, UT: Utah State U P, 2010. Print.
- American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). *Standards for Educational and Psychological Testing*. Washington, D.C.: AERA, 1999. Print.
- Angoff, William. "Validity: An Evolving Concept." *Test Validity*. Eds. Howard Wainer and Henry I. Bauer. Hillsdale, NJ: Lawrence Erlbaum Associates, 2009. 19-32. Print.
- Bakhtin, Mikhail. *The Dialogic Imagination: Four Essays*. 1975. Ed. Michael Holquist. Austin, TX: University of Texas Press, 1981. Print.
- Baldwin, Doug. "Fundamental Challenges in Developing and Scoring Constructed-Response Assessments." *Writing Assessment in the 21st Century*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012. 327-41. Print.
- Bamberg, Betty. "Conflicts Between Teaching and Assessing Writing: Using Program-Based Research to Resolve Pedagogical and Ethical Dilemmas." *The Writing Program Administrator as Researcher*. Eds. Shirley K. Rose and Irwin Weiser. Portsmouth, NH: Heinemann, 1999: 28-39. Print.
- Barshay, Jill. "New Common Core Exams Will Test Whether a Robo-Grader is as Accurate as a Human." *The Hechinger Report* [Columbia University, New York, NY] 25 Mar. 2014: Common Core. Web. 24 Mar. 2014.

- Behm, Nicholas N., Gregory R. Glau, Deborah H. Holdstein, Duan Roen, and Edward M. White., eds. *The WPA Outcomes Statement: A Decade Later*. Anderson, SC: Parlor Press, 2013. Print.
- Bedore, Pamela, and Deborah Rossen-Knill. "An Informed Self-Placement: Is a Choice Offered a Choice Received?" *Writing Program Administration* 28.1-2 (2004): 55-78. Print.
- Bennett, Randy Elliot, and William C. Ward, eds. *Construction Versus Choice in Cognitive Measurement: Issues in Constructed Response, Performance Testing, and Portfolio Assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993. Print.
- Bereiter, Carl. "Foreword." *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Eds. Mark D. Shermis and Jill C. Burstein. Mahwah, NJ: Lawrence Erlbaum, 2003. vii-x. Print.
- Berlak, Harold. "Race and the Achievement Gap." *Rethinking Multicultural Education: Teaching for Racial and Cultural Justice*. Ed. Wayne Au. Milwaukee, WI: Rethinking Schools, Ltd., 2009. 63-72. Print.
- Berlin, James. "Rhetoric, Poetic, and Culture: Contested Boundaries in English Studies." *The Politics of Writing Instruction: Postsecondary*. Eds. Richard Bullock, John Trimbur, and Charles Schuster. Portsmouth, NH: Boynton/Cook, 1991. 23-38. Print.

Blakesley, David, Erin J. Harvey, and Erica J. Reynolds. "Southern Illinois University Carbondale as an Institutional Model: The English 100/101 Stretch and Directed Self-Placement Program." *Directed Self-Placement: Principles and Practices*. Eds. Daniel J. Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. 31-47. Print.

Blakesley, David. "Directed Self-Placement in the University." *WPA: Writing Program Administration*, 25.2 (2002): 9-39. Print.

Breland, Hunter. Educational Testing Service. *The Direct Assessment of Writing Skill: A Measurement Review. College Board Report No. 83-6*. New York, NY: CEEB, 1983. Print.

Brereton, John. "A Closer Look at the Harvard Entrance Examinations in the 1870s." *Writing Assessment in the 21st Century*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012. 31-43. Print.

Broad, Bob. "Mapping a Dialectic with Edward M. White (In Four Scenes)." *Writing Assessment in the 21st Century*. Eds. Norbert Elliot and Les Perelman. New York, NY: Hampton Press, Inc., 2012. 259-69. Print.

---. *Organic Writing Assessment: Dynamic Criteria Mapping in Action*. Logan, UT: Utah State U P, 2009. Print.

---. *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Logan, UT: Utah State U P, 2003. Print.

Broome, Edwin Cornelius. *A Historical and Critical Discussion of College Admission Requirements*. 1903. New York: Columbia U P, 1963. Print.

- Brusin, Sandra. "Lessons Learned: Fits and Starts as We Try to Throw Open Those Gates." Conference on College Composition and Communication. JW Marriott, Indianapolis, IN. 22 Mar. 2014. Conference Presentation.
- Burstein, Jill. "Fostering Best Practices in Writing Assessment and Instruction with E-rater®." *Writing Assessment in the 21st Century*. Eds. Norbert Elliot and Les Perelman. New York, NY: Hampton Press, Inc., 2012. 203-17. Print.
- Burstein, Jill, and Daniel Marcu. "Automated Evaluation of Discourse Structure in Student Essays." *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Eds. Mark D. Shermis and Jill Burstein. Mahwah, NJ: Lawrence Erlbaum, 2003. 200-19. Print.
- Chernekoff, Janice. "Introducing Directed Self-Placement to Kutztown University." *Directed Self-Placement: Principles and Practices*. Eds. Daniel J. Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. 127-47. Print.
- Condon, William. "Large-Scale Assessment, Locally-Developed Measures, and Automated Scoring of Essays: Fishing for Red Herrings?" *Assessing Writing* 18.1 (2013): 100-08. Print.
- . "The Future of Portfolio-Based Writing Assessment: A Cautionary Tale." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012. 233-45. Print.
- Cooper, Peter. *The Assessment of Writing Ability: A Review of Research*. ETS Research Report 84-12. Princeton, NJ: Educational Testing Service, 1984. Print.

- Cornell, Cynthia E., and Robert D. Newton. "The Case of a Small Liberal Arts University." *Directed Self-Placement: Principles and Practices*. Eds. Daniel J. Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. 149-79. Print.
- Cronbach, Lee. "Test Validation." *Educational Measurement*. Ed. Robert Thorndike. Washington, D.C.: American Council on Education, 1971. 443-507. Print.
- Crowley, Sharon. *Toward a Civil Discourse: Rhetoric and Fundamentalism*. Pittsburgh, PA: University of Pittsburgh Press, 2006. Print.
- . "Response to Edward M. White." *Journal of Basic Writing* 15.1 (1996): 88-91. Print.
- Crusius, Timothy W. *Discourse: A Critique & Synthesis of Major Theories*. New York, NY: The Modern Languages Association of America, 1989. Print.
- Cushman, Ellen. "Editor's Introduction: Special Issue on Diversity and International Writing Assessment." *Research in the Teaching of English* 48.3 (2014): 269-70. Print.
- DasBender, Gita. "Assessing Generation 1.5 Learners: The Revelations of Directed Self-Placement." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012. 371-84. Print.
- Deane, Paul. "On the Relation between Automated Essay Scoring and Modern Views of the Writing Construct." *Assessing Writing* 18.1 (2013): 7-24. Print.
- Diederich, Paul B. *Measuring Growth in English*. Urbana, IL: NCTE, 1974. Print.
- Diederich, Paul B., John W. French, Sydel T. Carlton. *Factors in Judgments of Writing Ability. Research Bulletin RB-61-15*. Princeton, NJ: Educational Testing Service, 1961. Print.

- Dobrin, Sidney I. "Writing Takes Place." *Ecocomposition*. Eds. Christian R. Weissner and Sidney I. Dobrin. Albany: State University of New York Press, 2001. 11-25. Print.
- Elliot, Norbert. *On a Scale: A Social History of Writing Assessment in America*. New York: Peter Lang Publishing, 2005. Print.
- Elliot, Norbert, and David Williamson. "Automated Assessment of Writing." *Assessing Writing* 18.1 (2013): 1-6. Print.
- Elliot, Norbert, and Les Perelman, eds. *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. New York: Hampton Press, Inc., 2012. Print.
- Elliot, Norbert, Perry Deess, Alex Rudniy, and Kamal Joshi. "Placement of Students into First-Year Writing Courses." *Research in the Teaching of English* 46.3 (2012): 285-313. Print.
- Ericsson, Patricia Freitag, and Richard Haswell. *Machine Scoring of Student Essays: Truth and Consequences*. Logan, UT: Utah State UP, 2006. Print.
- "Engineering." Bureau of Labor Statistics. Occupational Outlook Handbook, 2010-11 Edition. U.S. Department of Labor. Web. 15 Mar. 2011.
- Faigley, Lester, Roger D. Cherry, David A. Jolliffe, and Anna M. Skinner. *Assessing Writers' Knowledge and Processes of Composing*. Norwood, NJ: Ablex, 1985. Print.
- Farley, Todd. *Making the Grades: My Misadventures in the Standardized Testing Industry*. San Francisco, CA: Berrett-Koehler Publishers, 2009. Print.

- Frus, Phyllis. "Directed Self-Placement at a Large Research University: A Writing Center Perspective." *Directed Self-Placement: Principles and Practices*. Eds. Daniel J. Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. 179-91. Print.
- Gallagher, Chris. "Being There: (Re)Making the Assessment Scene." *College Composition and Communication* 62.3 (2011): 450-76. Print.
- . *Reclaiming Assessment: A Better Alternative to the Accountability Agenda*. Portsmouth, NH: Heinemann, 2007. Print.
- Gere, Anne Ruggles, Laura Aull, Moisés Damián Perales Escudero, Zak Lancaster, and Elizabeth Vander Lei. "Local Assessment: Using Genre Analysis to Validate Directed Self-Placement." *College Composition and Communication* 64.4 (2013): 605-33. Print.
- Gere, Anne Ruggles, Laura Aull, Timothy Green, and Anne Porter. "Assessing the Validity of Directed Self-Placement at a Large University." *Assessing Writing* 15 (2010): 154-76. Print.
- Gilyard, Keith. *Voices of the Self*. Detroit, MI: Wayne State UP, 1991. Print.
- Glazer, Nancy. "Bringing the Test to the Teachers: Building a Bridge to a Standardized Writing Test." Conference on College Composition and Communication. The Riviera Hotel, Las Vegas, NV. 14 Mar. 2013. Conference Presentation.
- Godshalk, Fred I., Frances Swineford, and William E. Coffman. *The Measurement of Writing Ability*. New York, NY: CEEB, 1966. Print.
- "Grand Valley State University." *Wikipedia*. 11 Mar. 2014. Web. 29 Mar. 2014.
- Green, Robert L., and Robert J. Griffore. "The Impact of Standardized Testing on Minority Students." *The Journal of Negro Education* 49.3 (1980): 238-52.

- Guilford, Joy Paul. "New Standards for Test Evaluation." *Educational and Psychological Measurement* 6 (1946): 427-38. Print.
- Hall, Carla, and Jaffer Sheyholislami. "Using Appraisal Theory to Understand Rater Values: An Examination of Rater Comments on ESL Test Essays." *Journal of Writing Assessment* 6.1 (2013): n. pag. Web.
- Harrington, Susanmarie. "Learning to Ride the Waves: Making Decisions about Placement Testing." *WPA Journal* 28.3 (2005): 9-29. Print.
- Haswell, Richard. "Fighting Number with Number." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012: 413-24. Print.
- , ed. *Beyond Outcomes: Assessment and Instruction Within a University Writing Program*. Westport, Connecticut: Ablex, 2001. Print.
- . "Validation: Part of the Circle." *Beyond Outcomes: Assessment and Instruction Within a University Writing Program*. Ed. Richard Haswell. Westport, Connecticut: Ablex, 2001. 125-39. Print.
- . *Gaining Ground in College Writing: Tales of Development and Interpretation*. Dallas, TX: Southern Methodist University Press, 1991. Print.
- Herrington, Anne, and Charles Moran. "Writing to a Machine is Not Writing At All." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012: 219-32. Print.
- . "What Happens When Machines Read Our Students' Writing?" *College English* 63.4 (2001): 480-99. Print.

hooks, bell. *Teaching Community: A Pedagogy of Hope*. New York: Routledge, 2003.

Print.

Horner, Bruce, and Min-Zhan Lu. *Representing the "Other": Basic Writers and the*

Teaching of Basic Writing. Illinois: NCTE, 1999. Print.

Horner, Bruce, Min-Zhan Lu, Jacqueline Jones Royster, and John Trimbur. "Language

Difference in Writing: Toward a Translingual Approach." *College English* 73.3

(2011): 303-21. Print.

Huddleston, Edith M. "Measurement of Writing Ability at the College Entrance Level:

Objective vs. Subjective Testing Techniques." *Journal of Experimental*

Education, 22 (1954): 165-213. Print.

Hughes, Katherine, and Judith Scott-Clayton. Community College Resource Center.

Assessing Developmental Assessment in Community Colleges. CCRC Working

Paper No. 19. New York, NY: Columbia University, 2012. Print.

Hull, Glynda, Mike Rose, Kay Losey Fraser, and Marisa Castellano. "Remediation as

Social Construct: Perspectives from an Analysis of Classroom Discourse."

College Composition and Communication 42.3 (1991): 299-329. Print.

Huot, Brian. "Standing on His Shoulders: Understanding William L. Smith's

Contributions to Writing Assessment." *Blurring Boundaries: Developing Writers,*

Researchers, and Teachers. Ed. Peggy O'Neill. Cresskill, NJ: Hampton Press,

2007. 15-29. Print.

---. *(Re)Articulating Writing Assessment for Teaching and Learning*. Logan, UT: Utah

State U P, 2002. Print.

- . "Toward a New Theory of Writing Assessment." *College Composition and Communication* 47.4 (1996): 549-66. Print.
- Inoue, Asao B., and Mya Poe. *Race and Writing Assessment*. New York: Peter Lang Publishing, 2012. Print.
- Inoue, Asao B., Miriam Fernandez, Lemuel Gary, Mathew Gomes, Diana Harriger, Kristen Johnson, Sean Maddox, Brice Nakamura, Tyler Richmond, Aparna Sinha, and Daniel Speechly. "Directed Self-Placement, WPA-CompPile Research Bibliographies, No. 16." *WPA-CompPile Research Bibliographies*. 2011. Web.
- Jones, Edmund. "Self-Placement at a Distance: Challenge and Opportunities." *WPA Journal* 32.1 (2008): 57-75. Print.
- . "ACCUPLACER's Essay-Scoring Technology: When Reliability Does Not Equal Validity." *Machine Scoring of Student Essays: Truth and Consequences*. Eds. Patricia Freitag Ericsson and Richard Haswell. Logan, UT: Utah State UP, 2006. 93-113. Print.
- Kane, Michael. "Validating the Interpretations and Uses of Test Scores." *Journal of Educational Measurement*. 50.1 (2013): 1-73. Print.
- Karabel, Jerome. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston: Houghton Mifflin Harcourt, 2005. Print.
- Kent, Thomas, ed. *Post-Process Theory: Beyond the Writing-Process Paradigm*. Carbondale: Southern Illinois U P, 1999. Print.

- Ketai, Rachel Lewis. "Race, Remediation, and Readiness for College Writing: Reassessing the 'Self' in Directed Self-Placement." *Race and Writing Assessment*. Eds. Asao B. Inoue and Mya Poe. New York, NY: Peter Lang, 2012. 141-54. Print.
- Klobucar, Andrew, Paul Deane, Norbert Elliot, Chaitanya Ramineni, Perry Deess, and Alex Rudniy. "Automated Essay Scoring and the Search for Valid Writing Assessment." *International Advances in Writing Research: Cultures, Places, Measures*. Eds. Charles Bazerman, Chris Dean, Karen Lunsford, Susie Null, Paul Rogers, and Amanda Stansell. Fort Collins, CO: WAC Clearinghouse and Parlor Press, 2012. 103-19. Web.
- Lawrence, Ida, Gretchen W. Rigol, Thomas Van Essen, and Carol A. Jackson. *A Historical Perspective on the SAT 1926-2001. College Board Research Report No. 2002-7*. New York, NY: CEEB, 2002. Print.
- Lemann, Peter. *The Big Test*. New York, NY: Farrar, Straus and Giroux, 2000. Print.
- Lewiecki-Wilson, Cynthia, Jeff Sommers, and John Paul Tassoni. "Rhetoric and the Writer's Profile: Problematizing Directed Self-Placement." *Assessing Writing* 7.2 (2000), 165-83. Print.
- Lloyd-Jones, Richard. "Primary Trait Scoring." *Evaluating Writing: Describing, Measuring, Judging*. Eds. Charles R. Cooper and Lee Odell. Urbana, IL: NCTE, 1979. 33-66. Print.
- Lunsford, Andrea, and Lahoucine Ouzgane. *Crossing Borderlands: Composition and Postcolonial Studies*. Pittsburgh, PA: U of Pitt. P, 2004. Print.

Lynne, Patricia. *Coming to Terms: A Theory of Writing Assessment*. Logan, UT: Utah State U P, 2004. Print.

Messick, Samuel. "Standards of Validity and the Validity of Standards in Performance Assessment." *Educational Measurement: Issues and Practice* 14.4 (1995): 4-31. Print.

---. "Meaning and Value in Test Validation: The Science and Ethics of Assessment." *Educational Researcher* 18.2 (1989): 5-11. Print.

Migliaccio, Todd, and Dan Melzer. "Using Grounded Theory in Writing Assessment." *The WAC Journal* 22 (2011): 69-79. Print.

Moore, Cindy. "Changing the Language of Assessment: Lessons from Feminism." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Norbert Elliot and Les Perelman. New York: Hampton Press, Inc., 2012: 457-73. Print.

Moss, Pamela. "Testing the Test of the Test." *Assessing Writing* 5 (1998): 111-22. Print.

---. "Enlarging the Dialogue in Educational Measurement: Voices from Interpretative Research Traditions." *Educational Researcher* 25.1 (1996): 20-28. Print.

---. "Validity Without Reliability." *Educational Researcher* 23.2 (1994): 5-12. Print.

Murphy, Sandra. "At Last: Culture and Consequences: The Canaries in the Coal Mine." *Research in the Teaching of English* 42.2 (2007): 228-44. Print.

Nelson, Jennie, and Diane Kelly-Riley. "Students as Stakeholders: Maintaining a Responsive Assessment." *Beyond Outcomes: Assessment and Instruction Within a University Writing Program*. Ed. Richard Haswell. Westport, Connecticut: Ablex, 2001. 143-59. Print.

O'Neill, Peggy. "Reframing Reliability for Writing Assessment." *Journal of Writing Assessment* 4.1 (2011): n.p. Web.

O'Neill, Peggy, Cindy Moore, and Brian Huot. *Guide to College Writing Assessment*. Logan, UT: Utah State U P, 2009. Print.

Palmer, Orville. "Sixty Years of English Testing." *College Board Review* (1960): 8-14. Print.

Patelis, Thanos. "The Formation of the College Board and the Context of Today." *Annual Meeting of the National Council on Measurement in Education*. New Orleans, LA. 10 Apr. 2011. Presentation.

Peckham, Irvin. "Online Challenge versus Offline ACT." *College Composition and Communication* 61.4 (2010), 719-45. Print.

---. "Online Placement in First-Year Writing." *College Composition and Communication* 60.3 (2009), 517-40. Print.

Perelman, Les. "Critique of Mark D. Shermis & Ben Hammer, 'Contrasting State-of-the-Art Automated Scoring of Essays: Analysis.'" *Journal of Writing Assessment* 6 (2013): n. pag. Web.

---. "Construct Validity, Length, Score, and Time in Holistically Graded Writing Assessments: The Case Against Automated Essay Scoring (AEE)." *International Advances in Writing Research: Cultures, Places, Measures*. Eds. Charles Bazerman, Chris Dean, Karen Lunsford, Susie Null, Paul Rogers, and Amanda Stansell. Fort Collins, CO: WAC Clearinghouse and Parlor Press, 2012. 121-32. Web.

- . "Mass-Market Writing Assessments as Bullshit." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Les Perelman and Norbert Elliot. New York: Hampton Press, Inc., 2012. 425-37. Print.
- Petruzzi, Anthony. "Articulating a Hermeneutic Theory of Writing Assessment." *Assessing Writing* 13 (2008): 219-42. Print.
- Pinter, Robbie, and Ellen Sims. "Directed Self-Placement at Belmont University: Sharing Power, Forming Relationships, Fostering Reflection." *Directed Self-Placement: Principles and Practices*. Eds. Daniel J. Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. 107-125. Print.
- Poe, Mya. "The Consequences of Writing Assessment." *Research in the Teaching of English* 48.3 (2014): 271-75. Print.
- Poe, Mya, and Norbert Elliot. "Race and Writing Assessment: Cross-Disciplinary Frameworks for Impact Analysis." Conference on College Composition and Communication. Las Vegas, NV. 14 March 2013. Conference Presentation.
- "The Proper Use of Placement Test Scores." *University of Wisconsin System Center for Placement Testing*. Board of Regents of the University of Wisconsin System. 2011. 24 Mar. 2014.
- Ramineni, Chaitanya. "Validating Automated Essay Scoring for Online Writing Placement." *Assessing Writing* 18.1 (2013): 40-61. Print.
- Rose, Mike. "Colleges Need to Re-Mediate Remediation." *Mike Rose's Blog*. 28 Aug. 2009. Web. 26 Mar. 2014.
- . *Lives on the Boundary*. 1989. New York, NY: Penguin Books, 2005. Print.

- Royer, Daniel J., and Roger Gilles, eds. *Directed Self-Placement: Principles and Practices*. Cresskill, NJ: Hampton Press, 2003. Print.
- Schendel, Ellen, and Peggy O'Neill. "Exploring the Theories and Consequences of Self-Placement through Ethical Inquiry." *Assessing Writing* 6.2 (1999): 199-227. Print.
- Scott-Clayton, Judith. Community College Research Center. *Do High-Stakes Placement Exams Predict College Success? CCRC Working Paper No. 41*. New York, NY: Columbia University, 2012. Print.
- Shale, Doug. "Essay Reliability: Form and Meaning." *Assessment of Writing: Politics, Policies, Practices*. Eds. Edward M. White, William D. Lutz, and Sandra Kamusikiri. New York: MLA, 1996. 76-96. Print.
- Shaughnessy, Mina. *Errors and Expectations*. New York, NY: Oxford, 1977. Print.
- Shermis, Mark D., and Jill C. Burstein, eds. *Handbook of Automated Essay Evaluation*. New York, NY: Routledge, 2013. Print.
- . *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Mahwah, NJ: Lawrence Erlbaum, 2003. Print.
- Shermis, Mark D., Jill Burstein, and Sharon Apel Bursky. "Introduction to Automated Essay Evaluation." *Handbook of Automated Essay Evaluation*. Eds. Mark Shermis and Jill Burstein. New York, NY: Routledge, 2013. 1-15. Print.
- Shermis, Mark D., and Ben Hamner. *Contrasting State-of-the-Art Automated Scoring of Essays: Analysis*. 2012. pdf.
- "SIU Carbondale Enrollment History." *The Southern Illinoisian*. 6 May 2012. Web. 29 Mar. 2014.

- Smith, William. "The Importance of Teacher Knowledge in College Composition Placement Testing." *Reading Empirical Research Studies: The Rhetoric of Research*. Ed. R. J. Hayes. Norwood, NJ: Ablex, 1992. 289-316. Print.
- Stross, Randall. "The Algorithm Didn't Like My Essay." *The New York Times*. Business Section. 9 June 2012. Web. 24 Mar. 2014.
- "Students' Right to Their Own Language." *College Composition and Communication* 25.3 (1974): 1-18. Print.
- Thomson, Pat, and Inger Mewburn. "Why Do Academics Blog? It's Not for Public Outreach, Research Shows." *The Guardian* [London, England] 2 Dec. 2013: Higher Education Network. Web. 24 Mar. 2014.
- Trachsel, Mary. *Institutionalizing Literacy: The Historical Role of College Entrance Examinations in English*. Carbondale, IL: Southern Illinois University Press, 1992. Print.
- Vandenberg, Peter, Sue Hum, and Jennifer Clary-Lemon, eds. *Relations, Locations, Positions: Composition Theory for Writing Teachers*. Urbana, IL: NCTE, 2006. Print.
- Villanueva, Victor. *Bootstraps*. Urbana, IL: NCTE, 1993. Print.
- Walker, Paul. "Composition's Akrasia: The Devaluing of Intuitive Expertise in Writing Assessment." *Enculturation*. 15 (2013): n.pag. Web.
- White, Edward M. "My Five-Paragraph-Theme Theme." *College Composition and Communication* 59.3 (2008): 524-25. Print.
- . *Assigning, Responding, Evaluating*. Boston: Bedford/St. Martin's P, 2006. Print.

- . "The Misuse of Writing Assessment for Political Purposes." *Journal of Writing Assessment* 2.1 (2005): 21-36. Print.
- . "Foreword." *Directed Self-Placement: Principles and Practices*. Eds. Daniel Royer and Roger Gilles. Cresskill, NJ: Hampton Press, 2003. vii-viii. Print.
- . "Apologia for the Timed-Essay." *College Composition and Communication* 46 (1995): 30-45. Print.
- . "The Importance of Placement and Basic Studies." *Journal of Basic Writing* 14.2 (1995): 75-84. Print.
- . *Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance*. Hoboken, NJ: Jossey-Bass Inc., 1994. Print.
- White, Edward M., William D. Lutz, and Sandra Kamusikiri, eds. *Assessment of Writing: Politics, Policies, Practices*. New York: MLA, 1996. Print.
- Whithaus, Carl. "Foreword." *Handbook of Automated Essay Evaluation*. Eds. Mark Shermis and Jill Burstein. New York, NY: Routledge, 2013. vii-xi. Print.
- Williams, Juan. "Talk of the Nation." *National Public Radio*. 26 Feb. 2001. Radio.
- Williamson, Michael. "The Worship of Efficiency: Untangling Theoretical and Practical Considerations in Writing Assessment." *Assessing Writing* 1 (1994): 147-74.
- . "An Introduction to Holistic Scoring: The Social, Historical and Theoretical Context for Writing Assessment." *Validating Holistic Scoring for Writing Assessment*. Eds. Michael M. Williamson and Brian A. Huot. Cresskill, NJ: Hampton Press, 1993. Print. 1-43.

- Williamson, Michael, and Brian Huot, eds. *Validating Holistic Scoring for Writing Assessment: Theoretical and Empirical Foundations*. New York: Hampton Press, 1993. Print.
- Wilson, Maja. *Rethinking Rubrics in Writing Assessment*. Portsmouth, NH: Heinemann, 2006. Print.
- Winerip, Michael. "Facing a Robo-Grader? Just Keep Obfuscating Mellifluously." *The New York Times* 22 Apr. 2012: Education. Web. 24 Mar. 2014.
- "The World Factbook." *cia.gov*. Central Intelligence Agency. Web. 29 Mar. 2014.
- "Writing Assessment: A Position Statement." *Conference on College Composition and Communication*. 1995. Print.
- Yancey, Kathleen Blake. "The Rhetorical Situation of Writing Assessment: Exigence, Location, and the Making of Knowledge." *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White*. Eds. Les Perelman and Norbert Elliot. New York: Hampton Press, Inc., 2012. 475-92. Print.
- . "Looking Back as We Look Forward: Historicizing Writing Assessment." *College Composition and Communication* 50.3 (1999), 483-503. Print.

Appendix A

WPA Outcomes Statement for First-Year Composition

Adopted by the Council of Writing Program Administrators, April 2000; amended July 2008.

For further information about the development of the Outcomes Statement, please see <http://comppile.org/archives/WPAoutcomes/continue.html>

For further information about the Council of Writing Program Administrators, please see <http://www.wpacouncil.org>

Introduction

This statement describes the common knowledge, skills, and attitudes sought by first-year composition programs in American postsecondary education. To some extent, we seek to regularize what can be expected to be taught in first-year composition; to this end the document is not merely a compilation or summary of what currently takes place. Rather, the following statement articulates what composition teachers nationwide have learned from practice, research, and theory. This document intentionally defines only "outcomes," or types of results, and not "standards," or precise levels of achievement. The setting of standards should be left to specific institutions or specific groups of institutions.

Learning to write is a complex process, both individual and social, that takes place over time with continued practice and informed guidance. Therefore, it is important that teachers, administrators, and a concerned public do not imagine that these outcomes can be taught in reduced or simple ways. Helping students demonstrate these outcomes requires expert understanding of how students actually learn to write. For this reason we expect the primary audience for this document to be well-prepared college writing teachers and college writing program administrators. In some places, we have chosen to write in their professional language. Among such readers, terms such as "rhetorical" and "genre" convey a rich meaning that is not easily simplified. While we have also aimed at writing a document that the general public can understand, in limited cases we have aimed first at communicating effectively with expert writing teachers and writing program administrators.

These statements describe only what we expect to find at the end of first-year composition, at most schools a required general education course or sequence of courses. As writers move beyond first-year composition, their writing abilities do not merely improve. Rather, students' abilities not only diversify along disciplinary and professional lines but also move into whole new levels where expected outcomes expand, multiply, and diverge. For this reason, **each statement of outcomes for first-year composition is followed by suggestions for further work** that builds on these outcomes.

Rhetorical Knowledge

By the end of first year composition, students should

- Focus on a purpose
- Respond to the needs of different audiences
- Respond appropriately to different kinds of rhetorical situations
- Use conventions of format and structure appropriate to the rhetorical situation
- Adopt appropriate voice, tone, and level of formality
- Understand how genres shape reading and writing
- Write in several genres

Faculty in all programs and departments can build on this preparation by helping students learn

- The main features of writing in their fields
- The main uses of writing in their fields
- The expectations of readers in their fields

Critical Thinking, Reading, and Writing

By the end of first year composition, students should

- Use writing and reading for inquiry, learning, thinking, and communicating
- Understand a writing assignment as a series of tasks, including finding, evaluating, analyzing, and synthesizing appropriate primary and secondary sources
- Integrate their own ideas with those of others
- Understand the relationships among language, knowledge, and power

Faculty in all programs and departments can build on this preparation by helping students learn

- The uses of writing as a critical thinking method
- The interactions among critical thinking, critical reading, and writing
- The relationships among language, knowledge, and power in their fields

Processes

By the end of first year composition, students should

- Be aware that it usually takes multiple drafts to create and complete a successful text
- Develop flexible strategies for generating, revising, editing, and proof-reading
- Understand writing as an open process that permits writers to use later invention and re-thinking to revise their work
- Understand the collaborative and social aspects of writing processes
- Learn to critique their own and others' works
- Learn to balance the advantages of relying on others with the responsibility of doing their part
- Use a variety of technologies to address a range of audiences

Faculty in all programs and departments can build on this preparation by helping students learn

- To build final results in stages
- To review work-in-progress in collaborative peer groups for purposes other than editing
- To save extensive editing for later parts of the writing process
- To apply the technologies commonly used to research and communicate within their fields

Knowledge of Conventions

By the end of first year composition, students should

- Learn common formats for different kinds of texts
- Develop knowledge of genre conventions ranging from structure and paragraphing to tone and mechanics
- Practice appropriate means of documenting their work
- Control such surface features as syntax, grammar, punctuation, and spelling

Faculty in all programs and departments can build on this preparation by helping students learn

- The conventions of usage, specialized vocabulary, format, and documentation in their fields
- Strategies through which better control of conventions can be achieved

Composing in Electronic Environments

As has become clear over the last twenty years, writing in the 21st century involves the use of digital technologies for several purposes, from drafting to peer reviewing to editing. Therefore, although the *kinds* of composing processes and texts expected from students vary across programs and institutions, there are nonetheless common expectations.

By the end of first-year composition, students should:

- Use electronic environments for drafting, reviewing, revising, editing, and sharing texts
- Locate, evaluate, organize, and use research material collected from electronic sources, including scholarly library databases; other official databases (e.g., federal government databases); and informal electronic networks and internet sources
- Understand and exploit the differences in the rhetorical strategies and in the affordances available for both print and electronic composing processes and texts

Faculty in all programs and departments can build on this preparation by helping students learn

- How to engage in the electronic research and composing processes common in their fields
- How to disseminate texts in both print and electronic forms in their fields

Appendix B

CCCC Position Statement

A statement on an education issue approved by the CCCC Executive Committee

Writing Assessment: A Position Statement

Prepared by CCCC Committee on Assessment, November 2006 (revised March 2009)

Introduction

Writing assessment can be used for a variety of appropriate purposes, both inside the classroom and outside: providing assistance to students, awarding a grade, placing students in appropriate courses, allowing them to exit a course or sequence of courses, certifying proficiency, and evaluating programs—to name some of the more obvious. Given the high stakes nature of many of these assessment purposes, it is crucial that assessment practices be guided by sound principles to insure that they are valid, fair, and appropriate to the context and purposes for which they designed. This position statement aims to provide that guidance.

In spite of the diverse uses to which writing assessment is put, the general principles undergirding it are similar:

Assessments of written literacy should be designed and evaluated by well-informed current or future teachers of the students being assessed, for purposes clearly understood by all the participants; should elicit from student writers a variety of pieces, preferably over a substantial period of time; should encourage and reinforce good teaching practices; and should be solidly grounded in the latest research on language learning as well as accepted best assessment practices.

Guiding Principles for Assessment

1. Writing assessment is useful primarily as a means of improving teaching and learning. The primary purpose of any assessment should govern its design, its implementation, and the generation and dissemination of its results.

As a result...

A. Best assessment practice is informed by pedagogical and curricular goals, which are in turn formatively affected by the assessment. Teachers or administrators designing assessments should ground the assessment in the classroom, program or departmental context. The goals or outcomes assessed should lead to assessment data which is fed back to those involved with the regular activities assessed so that assessment results may be used to make changes in practice.

B. Best assessment practice is undertaken in response to local goals, not external pressures. Even when external forces require assessment, the local community must assert control of the assessment process, including selection of the assessment instrument and criteria.

C. Best assessment practice provides regular professional development opportunities. Colleges, universities, and secondary schools should make use of assessments as opportunities for professional development and for the exchange of information about student abilities and institutional expectations.

2. Writing is by definition social. Learning to write entails learning to accomplish a range of purposes for a range of audiences in a range of settings.

As a result...

A. Best assessment practice engages students in contextualized, meaningful writing. The assessment of writing must strive to set up writing tasks and situations that identify purposes appropriate to and appealing to the particular students being tested. Additionally, assessment must be contextualized in terms of why, where, and for what purpose it is being undertaken; this context must also be clear to the students being assessed and to all stakeholders.

B. Best assessment practice supports and harmonizes with what practice and research have demonstrated to be effective ways of teaching writing. What is easiest to measure—often by means of a multiple choice test—may correspond least to good writing; choosing a correct response from a set of possible answers is not composing. As

important, just asking students to write does not make the assessment instrument a good one. Essay tests that ask students to form and articulate opinions about some important issue, for instance, without time to reflect, talk to others, read on the subject, revise, and have a human audience promote distorted notions of what writing is. They also encourage poor teaching and little learning. Even teachers who recognize and employ the methods used by real writers in working with students can find their best efforts undercut by assessments such as these.

C. Best assessment practice is direct assessment by human readers. Assessment that isolates students and forbids discussion and feedback from others conflicts with what we know about language use and the benefits of social interaction during the writing process; it also is out of step with much classroom practice. Direct assessment in the classroom should provide response that serves formative purposes, helping writers develop and shape ideas, as well as organize, craft sentences, and edit. As stated by the CCCC Position Statement on Teaching, Learning, and Assessing Writing in Digital Environments, “we oppose the use of machine-scored writing in the assessment of writing.” Automated assessment programs do not respond as human readers. While they may promise consistency, they distort the very nature of writing as a complex and context-rich interaction between people. They simplify writing in ways that can mislead writers to focus more on structure and grammar than on what they are saying by using a given structure and style.

3. Any individual's writing ability is a sum of a variety of skills employed in a diversity of contexts, and individual ability fluctuates unevenly among these varieties.

As a result...

A. Best assessment practice uses multiple measures. One piece of writing—even if it is generated under the most desirable conditions—can never serve as an indicator of overall writing ability, particularly for high-stakes decisions. Ideally, writing ability must be assessed by more than one piece of writing, in more than one genre, written on different occasions, for different audiences, and responded to and evaluated by multiple readers as part of a substantial and sustained writing process.

B. Best assessment practice respects language variety and diversity and assesses writing on the basis of effectiveness for readers, acknowledging that as purposes vary, criteria will as well. Standardized tests that rely more on identifying grammatical and stylistic errors than authentic rhetorical choices disadvantage students whose home dialect is not the dominant dialect. Assessing authentic acts of writing simultaneously raises performance standards and provides multiple avenues to success. Thus students are not arbitrarily punished for linguistic differences that in some contexts make them more, not less, effective communicators. Furthermore, assessments that are keyed closely to an American cultural context may disadvantage second language writers. The CCCC Statement on Second Language Writing and Writers calls on us “to recognize the regular presence of second-language writers in writing classes, to understand their characteristics, and to develop instructional and administrative practices that are sensitive to their linguistic and cultural needs.” Best assessment practice responds to this call by creating assessments that are sensitive to the language varieties in use among the local population and sensitive to the context-specific outcomes being assessed.

C. Best assessment practice includes assessment by peers, instructors, and the student writer himself or herself. Valid assessment requires combining multiple perspectives on a performance and generating an overall assessment out of the combined descriptions of those multiple perspectives. As a result, assessments should include formative and summative assessments from all these kinds of readers. Reflection by the writer on her or his own writing processes and performances holds particular promise as a way of generating knowledge about writing and increasing the ability to write successfully.

4. Perceptions of writing are shaped by the methods and criteria used to assess writing.

As a result...

A. The methods and criteria that readers use to assess writing should be locally developed, deriving from the particular context and purposes for the writing being assessed. The individual writing program, institution, or consortium, should be recognized as a community of interpreters whose knowledge of context and purpose is integral to the assessment. There is no test which can be used in all environments for all purposes, and the best assessment for any group of students must be locally determined and may well be locally designed.

B. Best assessment practice clearly communicates what is valued and expected, and does not distort the nature of writing or writing practices. If ability to compose for various audiences is valued, then an assessment will assess this capability. For other contexts and purposes, other writing abilities might be valued, for instance, to develop a position on the basis of reading multiple sources or to compose a multi-media piece, using text and images. Values and purposes should drive assessment, not the reverse. A corollary to this statement is that assessment practices and criteria should change as conceptions of texts and values change.

C. Best assessment practice enables students to demonstrate what they do well in writing. Standardized tests tend to focus on readily accessed features of the language (grammatical correctness, stylistic choices) and on error rather than on the appropriateness of the rhetorical choices that have been made. Consequently, the outcome of such assessments is negative: students are said to demonstrate what they do wrong with language rather than what they do well. Quality assessments will provide the opportunity for students to demonstrate the ways they can write, displaying the strategies or skills taught in the relevant environment.

5. Assessment programs should be solidly grounded in the latest research on learning, writing, and assessment.

As a result...

A. Best assessment practice results from careful consideration of the costs and benefits of the range of available approaches. It may be tempting to choose an inexpensive, quick assessment, but decision-makers should consider the impact of assessment methods on students, faculty, and programs. The return on investment from the direct assessment of writing by instructor-evaluators includes student learning, professional development of faculty, and program development. These benefits far outweigh the presumed benefits of cost, speed, and simplicity that machine scoring might seem to promise.

B. Best assessment practice is continually under review and subject to change by well-informed faculty, administrators, and legislators. Anyone charged with the responsibility of designing an assessment program must be cognizant of the relevant research and must stay abreast of developments in the field. The theory and practice of writing assessment is continually informed by significant publications in professional journals and by presentations at regional and national conferences. The easy availability of this research to practitioners makes ignorance of its content reprehensible.

Applications to Assessment Settings

The guiding principles apply to assessment conducting in any setting. In addition, we offer the following guidelines for situations that may be encountered in specific settings.

Assessment in the Classroom

In a course context, writing assessment should be part of the highly social activity within the community of faculty and students in the class. This social activity includes:

- a period of ungraded work (prior to the completion of graded work) that receives response from multiple readers, including peer reviewers,
- assessment of texts—from initial through to final drafts—by human readers, and
- more than one opportunity to demonstrate outcomes.

Self-assessment should also be encouraged. Assessment practices and criteria should match the particular kind of text being created and its purpose. These criteria should be clearly communicated to students in advance so that the students can be guided by the criteria while writing.

Assessment for Placement

Placement criteria in the most responsible programs will be clearly connected to any differences in the available courses. Experienced instructor-evaluators can most effectively make a judgment regarding which course would best serve each student's needs and assign each student to the appropriate course. If scoring systems are used, scores should derive from criteria that grow out of the work of the courses into which students are being placed.

Decision-makers should carefully weigh the educational costs and benefits of timed tests, portfolios, directed self placement, etc. In the minds of those assessed, each of these methods implicitly establishes its value over that of others, so the first impact is likely to be on what students come to believe about writing. For example, timed writing may suggest to students that writing always cramps one for time and that real writing is always a test. Machine-scored tests may focus students on error-correction rather than on effective communication. In contrast, the value of portfolio assessment is that it honors the processes by which writers develop their ideas and re-negotiate how their communications are heard within a language community.

Students should have the right to weigh in on their assessment. Self-placement without direction may become merely a right to fail, whereas *directed* self-placement, either alone or in combination with other methods, provides not only useful

information but also involves and invests the student in making effective life decisions.

If for financial or even programmatic reasons the initial method of placement is somewhat reductive, instructors of record should create an opportunity early in the semester to review and change students' placement assignments, and uniform procedures should be established to facilitate the easy re-placement of improperly placed students. Even when the placement process entails direct assessment of writing, the system should accommodate the possibility of improper placement. If assessment employs machine scoring, whether of actual writing or of items designed to elicit error, it is particularly essential that every effort be made through statistical verification to see that students, individually and collectively, are placed in courses that can appropriately address their skills and abilities.

Placement processes should be continually assessed and revised in accord with course content and overall program goals. This is especially important when machine-scored assessments are used. Using methods that are employed uniformly, teachers of record should verify that students are appropriately placed. If students are placed according to scores on such tests, the ranges of placement must be revisited regularly to accommodate changes in curricula and shifts in the abilities of the student population.

Assessment of Proficiency

Proficiency or exit assessment involves high stakes for students. In this context, assessments that make use of substantial and sustained writing processes are especially important.

Judgments of proficiency must also be made on the basis of performances in multiple and varied writing situations (for example, a variety of topics, audiences, purposes, genres).

The assessment criteria should be clearly connected to desired outcomes. When proficiency is being determined, the assessment should be informed by such things as the core abilities adopted by the institution, the course outcomes established for a program, and/or the stated outcomes of a single course or class. Assessments that do not address such outcomes lack validity in determining proficiency.

The higher the stakes, the more important it is that assessment be direct rather than indirect, based on actual writing rather than on answers on multiple-choice tests, and evaluated by people involved in the instruction of the student rather than via machine scoring. To evaluate the proficiency of a writer on other criteria than multiple writing tasks and situations is essentially disrespectful of the writer.

Assessment of Programs

Program assessment refers to evaluations of performance in a large group, such as students in a multi-section course or majors graduating from a department. Because assessment offers information about student performance and the factors which affect that performance, it is an important way for programs or departments to monitor and develop their practice.

Programs and departments should see themselves as communities of professionals whose assessment activities reveal common values, provide opportunities for inquiry and debate about unsettled issues, and communicate measures of effectiveness to those inside and outside the program. Members of the community are in the best position to guide decisions about what assessments will best inform that community. It is important to bear in mind that random sampling of students can often provide large-scale information and that regular assessment should affect practice.

Assessment for School Admission

Admissions tests are not only high stakes for students, they are also an extremely important component for educational institutions determining if they and a student are an appropriate match. Consequently, where students' writing ability is a factor in the admissions decision, the writing assessments should consist of direct measures of actual writing. Moreover, the assessment should consist of multiple writing tasks and should allow sufficient time for a student to engage in all stages of the writing process.

Assessments should be appropriate to educational institutions' distinctive missions and student populations, although similar institutions may collaborate to create assessments. Assessment should be developed in consultation with high school writing teachers.

This position statement may be printed, copied, and disseminated without permission from NCTE.

Appendix C



CRITICAL THINKING VALUE RUBRIC

for more information, please contact rubric@aaau.org

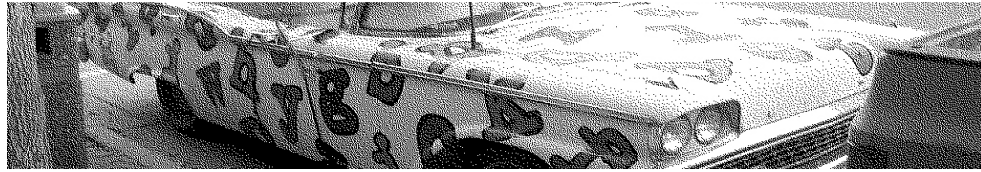
Definition

Critical thinking is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion.

Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (all one) level performance.

	Capstone 4	Milestones 3 2		Benchmark 1
Explanation of issues	Issue/ problem to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding.	Issue/ problem to be considered critically is stated, described, and clarified so that understanding is not seriously impeded by omissions.	Issue/ problem to be considered critically is stated but description leaves some terms undefined, ambiguities unexplored, boundaries undetermined, and/ or backgrounds unknown.	Issue/ problem to be considered critically is stated without clarification or description.
Evidence <i>Selecting and using information to investigate a point of view or conclusion</i>	Information is taken from sources(s) with enough interpretation/ evaluation to develop a comprehensive analysis or synthesis. Viewpoints of experts are questioned thoroughly.	Information is taken from sources(s) with enough interpretation/ evaluation to develop a coherent analysis or synthesis. Viewpoints of experts are subject to questioning.	Information is taken from sources(s) with some interpretation/ evaluation, but not enough to develop a coherent analysis or synthesis. Viewpoints of experts are taken as mostly fact, with little questioning.	Information is taken from sources(s) without any interpretation/ evaluation. Viewpoints of experts are taken as fact, without question.
Influence of context and assumptions	Thoroughly (systematically and methodically) analyzes own and others' assumptions and carefully evaluates the relevance of contexts when presenting a position.	Identifies own and others' assumptions and several relevant contexts when presenting a position.	Questions some assumptions. Identifies several relevant contexts when presenting a position. May be more aware of others' assumptions than one's own (or vice versa).	Shows an emerging awareness of present assumptions (sometimes labels assertions as assumptions). Begins to identify some contexts when presenting a position.
Student's position (perspective, thesis/ hypothesis)	Specific position (perspective, thesis/ hypothesis) is imaginative, taking into account the complexities of an issue. Limits of position (perspective, thesis/ hypothesis) are acknowledged. Others' points of view are synthesized within position (perspective, thesis/ hypothesis).	Specific position (perspective, thesis/ hypothesis) takes into account the complexities of an issue. Others' points of view are acknowledged within position (perspective, thesis/ hypothesis).	Specific position (perspective, thesis/ hypothesis) acknowledges different sides of an issue.	Specific position (perspective, thesis/ hypothesis) is stated, but is simplistic and obvious.
Conclusions and related outcomes (implications and consequences)	Conclusions and related outcomes (consequences and implications) are logical and reflect student's informed evaluation and ability to place evidence and perspectives discussed in priority order.	Conclusion is logically tied to a range of information, including opposing viewpoints. Related outcomes (consequences and implications) are identified clearly.	Conclusion is logically tied to information (because information is chosen to fit the desired conclusion); some related outcomes (consequences and implications) are identified clearly.	Conclusion is inconsistently tied to some of the information discussed; related outcomes (consequences and implications) are oversimplified.

Appendix D



GOALS AND OUTCOMES FOR PORTFOLIO ASSESSMENT IN ENGLISH 095

In the INTERPRETIVE ESSAY:

- Maintain a controlling purpose that...
 - Reflects your shared concerns with others.
 - Reflects what matters for others who are addressed in and affected by your interpretation.
 - Creates coherence throughout the essay.
- Critically interpret course texts to support your purpose by...
 - Going beyond summary.
 - Incorporating and examining multiple passages.
 - Identifying and analyzing the language choices writers make.
 - Identifying and analyzing the composing decisions writers make.
- Demonstrate awareness of your rhetorical situation by...
 - Providing relevant context to readers through the use of background information, brief summaries, definitions of key terms, and examples.
 - Making consistent language choices and composing decisions appropriate to your purpose, context, and audience.

In the REFLECTIVE ESSAY, account for and evaluate...

- The language and design choices you made in your interpretive essay.
- The reading, writing, and revision practices you undertook throughout the semester to create your academic text.
- Your experiences in the course and their impact on how you created various texts.

GOALS AND OUTCOMES FOR PORTFOLIO ASSESSMENT IN ENGLISH 101 AND ESL 118

With the INTERPRETIVE ESSAYS:

- Maintain a controlling purpose that...
 - reflects what matters to or is at stake for you in the interpretation.
 - responds to what matters to or is at stake for those you address and those you respond to in the writing.
 - creates and maintains coherence and clarity for readers by shaping patterns of arrangement and other writing choices.
- Critically interpret course texts by...
 - incorporating, contextualizing, and examining multiple passages from the texts being interpreted.
 - identifying and analyzing strategies and choices, including key terms, distinctions and questions being asked within the texts being interpreted.
 - describing and evaluating how the arrangement, design choices, and other material conditions of the interpreted texts shape a reader's understanding.
- Make use of academic conventions for writing by...
 - providing relevant context of the texts being interpreted for readers, such as background information, brief summaries, definitions of key terms, and examples.
 - integrating and documenting all summaries, paraphrases, and direct quotations accurately and fairly and by following current MLA guidelines.
 - demonstrating an ability to meet academic expectations for grammar and mechanics in final drafts.

Present REFLECTIVE WRITING that accounts for and evaluates the choices made in the interpretive essays by...

- considering the composing and design strategies you developed through successive revisions given your rhetorical situation.
- describing how your writing responds ethically to what matters or is at stake for others who are addressed or affected by your interpretive writing.

GOALS AND OUTCOMES FOR PORTFOLIO ASSESSMENT IN ENGLISH 102

In the portfolio, you will:

- Present AN INQUIRY-BASED RESEARCH PROJECT that...
 - has a clearly defined central research question that reflects your concerns and interests.
 - develops knowledge, insight, or perspective about the matter being researched by engaging in critical inquiry.
 - makes appropriate use of sources, including scholarly sources.
- Maintain a controlling purpose that...
 - emerges from the central research question.
 - responds ethically to what matters or is at stake for others who are addressed or affected by the research project.
 - creates and maintains coherence and clarity for readers by shaping patterns of arrangement, design, and other writing choices.
- Engage in critical inquiry in ways that support your purpose by...
 - going beyond summary to position yourself and your ideas in relation to the ideas of others by engaging sources through interpretation, analysis, or critique.
 - using sources to frame or critically question other sources or issues.
 - describing and analyzing the positioning of sources in relation to each other and why those contexts matter.
- Follow writing conventions appropriate to the rhetorical situation of the writing, including...
 - providing relevant context to readers such as background information, examples, definitions, etc.
 - integrating the ideas of others accurately and fairly through summary, paraphrase, and direct quotation.
 - documenting all sources with in-text citations and a works cited or references page following current MLA, APA, or Chicago Manual of Style guidelines, as specified by the instructor.
 - demonstrating an ability to meet academic expectations for grammar and mechanics.
- Present REFLECTIVE WRITING that accounts for and evaluates...
 - the evolution of your controlling purpose, through successive revisions, in relation to the project's audience(s), stakeholders, and your own interests
 - your composing and design strategies, developed through successive revisions, given your rhetorical situation
 - how and why sources were chosen and how you used them in the project
 - your inquiry-based project in relation to your understanding of academic research and writing

Appendix E

2012 AP[®] ENGLISH LANGUAGE AND COMPOSITION FREE-RESPONSE QUESTIONS

Question 3

(Suggested time—40 minutes. This question counts for one-third of the total essay section score.)

Consider the distinct perspectives expressed in the following statements.

If you develop the absolute sense of certainty that powerful beliefs provide, then you can get yourself to accomplish virtually anything, including those things that other people are certain are impossible.

William Lyon Phelps, American educator, journalist, and professor (1865–1943)

I think we ought always to entertain our opinions with some measure of doubt. I shouldn't wish people dogmatically to believe any philosophy, not even mine.

Bertrand Russell, British author, mathematician, and philosopher (1872–1970)

In a well-organized essay, take a position on the relationship between certainty and doubt. Support your argument with appropriate evidence and examples.

AP[®] ENGLISH LANGUAGE AND COMPOSITION 2012 SCORING GUIDELINES

Question 3

The score should reflect a judgment of the essay's quality as a whole. Remember that students had only 40 minutes to read and write; the essay, therefore, is not a finished product and should not be judged by standards appropriate for an out-of-class assignment. Evaluate the essay as a draft, making certain to reward students for what they do well.

All essays, even those scored 8 or 9, may contain occasional lapses in analysis, prose style, or mechanics. Such features should enter into the holistic evaluation of an essay's overall quality. In no case may an essay with many distracting errors in grammar and mechanics be scored higher than a 2.

- 9** Essays earning a score of 9 meet the criteria for a score of 8 and, in addition, are especially sophisticated in their argument, thorough in their development, or particularly impressive in their control of language.

8 Effective

Essays earning a score of 8 **effectively** develop a position on the relationship between certainty and doubt. The evidence and explanations used are appropriate and convincing, and the argument is especially coherent and well developed. The prose demonstrates a consistent ability to control a wide range of the elements of effective writing but is not necessarily flawless.

- 7** Essays earning a score of 7 meet the criteria for a score of 6 but provide a more complete explanation, more thorough development, or a more mature prose style.

6 Adequate

Essays earning a score of 6 **adequately** develop a position on the relationship between certainty and doubt. The evidence and explanations used are appropriate and sufficient, and the argument is coherent and adequately developed. The writing may contain lapses in diction or syntax, but generally the prose is clear.

- 5** Essays earning a score of 5 develop a position on the relationship between certainty and doubt. The evidence or explanations used may be uneven, inconsistent, or limited. The writing may contain lapses in diction or syntax, but it usually conveys the student's ideas.

4 Inadequate

Essays earning a score of 4 **inadequately** develop a position on the relationship between certainty and doubt. The evidence or explanations used may be inappropriate, insufficient, or less convincing. The argument may have lapses in coherence or be inadequately developed. The prose generally conveys the student's ideas but may be less consistent in controlling the elements of effective writing.

- 3** Essays earning a score of 3 meet the criteria for a score of 4 but demonstrate less success in developing a position on the relationship between certainty and doubt. The essays may show less maturity in control of writing.

**AP[®] ENGLISH LANGUAGE AND COMPOSITION
2012 SCORING GUIDELINES**

Question 3 (continued)

2 Little Success

Essays earning a score of 2 demonstrate **little success** in developing a position on the relationship between certainty and doubt. These essays may misunderstand the prompt or substitute a simpler task by responding to the prompt tangentially with unrelated, inaccurate, or inappropriate explanation. The prose often demonstrates consistent weaknesses in writing, such as grammatical problems, a lack of development or organization, or a lack of coherence and control.

- 1** Essays earning a score of 1 meet the criteria for a score of 2 but are undeveloped, especially simplistic in their explanation and argument, weak in their control of language, or especially lacking in coherence and development.
- 0** Indicates an off-topic response, one that merely repeats the prompt, an entirely crossed-out response, a drawing, or a response in a language other than English.
- Indicates an entirely blank response.

Appendix F



National Council of Teachers of English

A Professional Association of Educators in English Studies, Literacy, and Language Arts

NCTE Position Statement on Machine Scoring

Machine Scoring Fails the Test

Approved by the NCTE Executive Committee, April 2013

[A] computer could not measure accuracy, reasoning, adequacy of evidence, good sense, ethical stance, convincing argument, meaningful organization, clarity, and veracity in your essay. If this is true I don't believe a computer would be able to measure my full capabilities and grade me fairly. -- **Akash, student**

[H]ow can the feedback a computer gives match the carefully considered comments a teacher leaves in the margins or at the end of your paper? -- **Pinar, student**

(Responses to *New York Times* The Learning Network blog post, "How Would You Feel about a Computer Grading Your Essays?", 5 April 2013)

Writing is a highly complex ability developed over years of practice, across a wide range of tasks and contexts, and with copious, meaningful feedback. Students must have this kind of sustained experience to meet the demands of higher education, the needs of a 21st-century workforce, the challenges of civic participation, and the realization of full, meaningful lives.

As the Common Core State Standards (CCSS) sweep into individual classrooms, they bring with them a renewed sense of the importance of writing to students' education. Writing teachers have found many aspects of the CCSS to applaud; however, we must be diligent in developing assessment systems that do not threaten the possibilities for the rich, multifaceted approach to writing instruction advocated in the CCSS. Effective writing assessments need to account for the nature of writing, the ways students develop writing ability, and the role of the teacher in fostering that development.

Research¹ on the assessment of student writing consistently shows that high-stakes writing tests alter the normal conditions of writing by denying students the opportunity to think, read, talk with others, address real audiences, develop ideas, and revise their emerging texts over time. Often, the results of such tests can affect the livelihoods of teachers, the fate of schools, or the educational opportunities for students. In such conditions, the narrowly conceived, artificial form of the tests begins to subvert attention to other purposes and varieties of writing development in the classroom. Eventually, the tests erode the foundations of excellence in writing instruction, resulting in students who are less prepared to meet the demands of their continued education and future occupations. Especially in the transition from high school to college, students are ill served when their writing experience has been dictated by tests that ignore the ever-more complex and varied types and uses of writing found in higher education.

Note: (1) All references to research are supported by the extensive work documented in the annotated bibliography attached to this report. The bibliography is drawn from a body of independent and industry research that supports other critiques of machine scoring, such as the [Professionals Against Machine Scoring Of Student Essays In High-Stakes Assessment Petition Initiative](http://humanreaders.org/petition/) (<http://humanreaders.org/petition/>).

These concerns -- increasingly voiced by parents, teachers, school administrators, students, and members of the general public -- are intensified by the use of machine-scoring systems to read and evaluate students' writing. To meet the outcomes of the Common Core State Standards, various consortia, private corporations, and testing agencies propose to use computerized assessments of student writing. The attraction is obvious: once programmed, machines might reduce the costs otherwise associated with the human labor of reading, interpreting, and evaluating the writing of our students. Yet when we consider what is lost because of machine scoring, the presumed savings turn into significant new costs -- to students, to our educational institutions, and to society. Here's why:

- Computers are unable to recognize or judge those elements that we most associate with good writing (logic, clarity, accuracy, ideas relevant to a specific topic, innovative style, effective appeals to audience, different forms of organization, types of persuasion, quality of evidence, humor or irony, and effective uses of repetition, to name just a few). Using computers to "read" and evaluate students' writing (1) denies students the chance to have anything but limited features recognized in their writing; and (2) compels teachers to ignore what is most important in writing instruction in order to teach what is least important.
- Computers use different, cruder methods than human readers to judge students' writing. For example, some systems

gauge the sophistication of vocabulary by measuring the average length of words and how often the words are used in a corpus of texts; or they gauge the development of ideas by counting the length and number of sentences per paragraph.

- Computers are programmed to score papers written to very specific prompts, reducing the incentive for teachers to develop innovative and creative occasions for writing, even for assessment.
- Computers get progressively worse at scoring as the length of the writing increases, compelling test makers to design shorter writing tasks that don't represent the range and variety of writing assignments needed to prepare students for the more complex writing they will encounter in college.
- Computer scoring favors the most objective, "surface" features of writing (grammar, spelling, punctuation), but problems in these areas are often created by the testing conditions and are the most easily rectified in normal writing conditions when there is time to revise and edit. Privileging surface features disproportionately penalizes nonnative speakers of English who may be on a developmental path that machine scoring fails to recognize.
- Conclusions that computers can score as well as humans are the result of humans being trained to score like the computers (for example, being told not to make judgments on the accuracy of information).
- Computer scoring systems can be "gamed" because they are poor at working with human language, further weakening the validity of their assessments and separating students not on the basis of writing ability but on whether they know and can use machine-tricking strategies.
- Computer scoring discriminates against students who are less familiar with using technology to write or complete tests. Further, machine scoring disadvantages school districts that lack funds to provide technology tools for every student and skews technology acquisition toward devices needed to meet testing requirements.
- Computer scoring removes the purpose from written communication -- to create human interactions through a complex, socially consequential system of meaning making -- and sends a message to students that writing is not worth their time because reading it is not worth the time of the people teaching and assessing them.

What Are the Alternatives?

Together with other professional organizations, the National Council of Teachers of English has established research-based guidelines for effective teaching and assessment of writing, such as the *Standards for the Assessment of Reading and Writing* (rev. ed., 2009), the *Framework for Success in Postsecondary Writing* (2011), the *NCTE Beliefs about the Teaching of Writing* (2004), and the *Framework for 21st Century Curriculum and Assessment* (2008, 2013). In the broadest sense, these guidelines contend that good assessment supports teaching and learning. Specifically, high-quality assessment practices will

- encourage students to become engaged in literacy learning, to reflect on their own reading and writing in productive ways, and to set respective literacy goals;
- yield high-quality, useful information to inform teachers about curriculum, instruction, and the assessment process itself;
- balance the need to assess summatively (make final judgments about the quality of student work) with the need to assess formatively (engage in ongoing, in-process judgments about what students know and can do, and what to teach next);
- recognize the complexity of literacy in today's society and reflect that richness through holistic, authentic, and varied writing instruction;
- at their core, involve professionals who are experienced in teaching writing, knowledgeable about students' literacy development, and familiar with current research in literacy education.

A number of effective practices enact these research-based principles, including portfolio assessment; teacher assessment teams; balanced assessment plans that involve more localized (classroom- and district-based) assessments designed and administered by classroom teachers; and "audit" teams of teachers, teacher educators, and writing specialists who visit districts to review samples of student work and the curriculum that has yielded them. We focus briefly here on portfolios because of the extensive scholarship that supports them and the positive experience that many educators, schools, and school districts have had with them.

Engaging teams of teachers in evaluating portfolios at the building, district, or state level has the potential to honor the challenging expectations of the CCSS while also reflecting what we know about effective assessment practices. Portfolios offer the opportunity to

- look at student writing across multiple events, capturing growth over time while avoiding the limitations of "one test on one day";
- look at the range of writing across a group of students while preserving the individual character of each student's writing;
- review student writing through multiple lenses, including content accuracy and use of resources;
- assess student writing in the context of local values and goals as well as national standards.

Just as portfolios provide multiple types of data for assessment, they also allow students to *learn* as a result of engaging in the assessment process, something seldom associated with more traditional one-time assessments. Students gain insight about their own writing, about ways to identify and describe its growth, and about how others – human readers – interpret their work. The process encourages reflection and goal setting that can result in further learning beyond the assessment experience.

Similarly, teachers grow as a result of administering and scoring the portfolio assessments, something seldom associated with more traditional one-time assessments. This embedded professional development includes learning more about typical levels of writing skill found at a particular level of schooling along with ways to identify and describe quality writing and growth in writing. The discussions about collections of writing samples and criteria for assessing the writing contribute to a shared investment among all participating teachers in the writing growth of all students. Further, when the portfolios include a wide range of artifacts from learning and writing experiences, teachers assessing the portfolios learn new ideas for classroom instruction as well as ways to design more sophisticated methods of assessing student work on a daily basis.

Several states such as Kentucky, Nebraska, Vermont, and California have experimented with the development of large-scale portfolio assessment projects that make use of teams of teachers working collaboratively to assess samples of student work. Rather than investing heavily in assessment plans that cannot meet the goals of the CCSS, various legislative groups, private companies, and educational institutions could direct those funds into refining these nascent portfolio assessment systems. This investment would also support teacher professional development and enhance the quality of instruction in classrooms – something that machine-scored writing prompts cannot offer.

What's Next

In 2010, the federal government awarded \$330 million to two consortia of states "to provide ongoing feedback to teachers during the course of the school year, measure annual school growth, and move beyond narrowly focused bubble tests" (United States Department of Education). Further, these assessments will need to align to the new standards for learning in English and mathematics. This has proven to be a formidable task, but it is achievable. By combining the already existing National Assessment of Educational Progress (NAEP) assessment structures for evaluating school system performance with ongoing portfolio assessment of student learning by educators, we can cost-effectively assess writing without relying on flawed machine-scoring methods. By doing so, we can simultaneously deepen student and educator learning while promoting grass-roots innovation at the classroom level. For a fraction of the cost in time and money of building a new generation of machine assessments, we can invest in rigorous assessment and teaching processes that enrich, rather than interrupt, high-quality instruction. Our students and their families deserve it, the research base supports it, and literacy educators and administrators will welcome it.

Work Cited

United States Department of Education. "[U.S. Secretary of Education Duncan Announces Winners of Competition to Improve Student Assessments](http://www.ed.gov/news/press-releases/us-secretary-education-duncan-announces-winners-competition-improve-student-assessments) [<http://www.ed.gov/news/press-releases/us-secretary-education-duncan-announces-winners-competition-improve-student-assessments>]." (2 Sept 2010; retrieved 11 April 2013).

Appendix G

Executive Summary

The concept of “college readiness” is increasingly important in discussions about students’ preparation for postsecondary education.

This Framework describes the rhetorical and twenty-first-century skills as well as habits of mind and experiences that are critical for college success. Based in current research in writing and writing pedagogy, the Framework was written and reviewed by two- and four-year college and high school writing faculty nationwide and is endorsed by the Council of Writing Program Administrators, the National Council of Teachers of English, and the National Writing Project.

Habits of mind refers to ways of approaching learning that are both intellectual and practical and that will support students’ success in a variety of fields and disciplines. The Framework identifies eight habits of mind essential for success in college writing:

- Curiosity – the desire to know more about the world.
- Openness – the willingness to consider new ways of being and thinking in the world.
- Engagement – a sense of investment and involvement in learning.
- Creativity – the ability to use novel approaches for generating, investigating, and representing ideas.
- Persistence – the ability to sustain interest in and attention to short- and long-term projects.
- Responsibility – the ability to take ownership of one’s actions and understand the consequences of those actions for oneself and others.
- Flexibility – the ability to adapt to situations, expectations, or demands.
- Metacognition – the ability to reflect on one’s own thinking as well as on the individual and cultural processes used to structure knowledge.

The Framework then explains how teachers can foster these habits of mind through **writing, reading, and critical analysis** experiences. These experiences aim to develop students’

- Rhetorical knowledge – the ability to analyze and act on understandings of audiences, purposes, and contexts in creating and comprehending texts;
- Critical thinking – the ability to analyze a situation or text and make thoughtful decisions based on that analysis, through writing, reading, and research;
- Writing processes – multiple strategies to approach and undertake writing and research;
- Knowledge of conventions – the formal and informal guidelines that define what is considered to be correct and appropriate, or incorrect and inappropriate, in a piece of writing; and
- Ability to compose in multiple environments – from traditional pen and paper to electronic technologies.

CURRICULUM VITAE FOR JESSICA NASTAL-DEMA

Department of English
University of Wisconsin—Milwaukee
Milwaukee, WI 53201
jlnastal@uwm.edu

EDUCATION

Ph.D. English: Rhetoric and Composition, 2014

University of Wisconsin-Milwaukee

Dissertation: *(Re)Envisioning Placement for 21st Century Writing Programs*

Director: Charles Schuster, PhD

M.A. English, 2007

Saint Louis University

Graduated with distinction

M.A. English Cultural and Literary Studies, 2007

Universidad Autónoma de Madrid

B.A. English; Political Science minor, 2002

Loyola University Chicago

Graduated magna cum laude and Phi Beta Kappa

PROFESSIONAL EXPERIENCE

University of Wisconsin-Milwaukee

Editor, WiscAMP Program, 2013

Lecturer, Intensive English Program, 2010-2013

Graduate Teaching Assistant, instructor of record, 2009-2013

Saint Louis University

Graduate Teaching Assistant, instructor of record, 2005-07

Writing Center Instructor, 2005-06

HONORS & AWARDS

Appointed as Associate Editor, *Journal of Writing Assessment*, 2011-present

University Distinguished Dissertation Fellowship, \$16,500, 2013-14

Chancellor's Incoming Graduate Student Award, \$12,500, 2009-10

Chancellor's Graduate Student Award, 2012, 2013

English Department Teaching Excellence Award, 2012

Travel Award, Distinguished Dissertation Fellow, UWM, 2013-14

Travel Grant, Council of Writing Program Administrators, 2013

Travel Award, Graduate School, UWM, 2012, 2013

Travel Award, English Department, UWM, 2010, 2011, 2012, 2014

COURSES TAUGHT

University of Wisconsin—Milwaukee

Composition

ENGL 095, Fundamentals of College Reading and Writing
 ENGL 101, Introduction to College Writing
 ENGL 102, College Writing and Research
 ESL 118, Advanced College Writing in English as a Second Language
 ENGL 102, College Writing and Research for ESL students

Intensive English Program

Reading
 Writing
 Electives, including: Cinema, The Short Story, Alternative Literature,
 Current Events, Writing for Beginners, and TOEFL Preparation
 Academic Reading and Writing
 Cultural Communication
 Reading and Writing for students enrolled dually in the IEP and university

200-Level

ENGL 215, Introduction to English Studies

Saint Louis University (Madrid Campus)

Composition

ENGL 150, The Process of Writing
 ENGL 190, Advanced Strategies of Rhetoric and Research
 ENGL 192, Advanced Writing for Professionals

Intensive English

Reading
 Advanced Listening, Grammar, and Reading

PUBLICATIONS

- Nastal, Jessica. "Review of: D.21 Race and Writing Assessment: Cross-Disciplinary Frameworks for Impact Analysis." *Kairos* 18.1 (2013): n. pag. Web.
- Nastal, Jessica. "Review of: *Writing Assessment in the 21st Century: Essays in Honor of Edward M. White.*" *Journal of Writing Assessment Reading List* (2012, 2013): n. pag. Web. <http://jwareadinglist.blogspot.com/>
- Nastal, Jessica and Ryan Witt. "Review of 'Understanding Job Advertisements and Composing Effective Cover Letters I & II' Presentations at CWPA 2012." *CWPA Mentor Blog*. (2012). Web.
- Nastal, Jessica. "The Tyrone Family as Allegory for pre-World War Two United States of America: Illusions of an Isolationist Nation." *eO'Neill.com*. (2010). Web.
- Nastal, Jessica. "Madrid." *Mosaic Literary Magazine*. 5 (2009). Print.

CONFERENCE PRESENTATIONS

- “Enacting a Responsive Writing Program.” Panel Participant at *Watson Conference*. 2014.
- “In Search of Assessments that Work.” Panel participant at *CWPA*. 2014
- “Critical Junctures: Exploring Different Sites of Learning.” Panel chairperson at *CCCC*. 2014.
- Editors’ Roundtable participant. Research Network Forum at *CCCC*. 2014.
- “(Re)Envisioning Placement in the 21st Century.” (accepted) *North Carolina Symposium on Teaching Writing*. 2014.
- “Writing Assessment as Enculturation into the University.” (accepted) *Making Meaning: Language, Rhetoric, and Enculturation Conference at UM*. 2013.
- “Traveling On the Feedback Loop: Using Writing Program Assessment to Strengthen Undergraduate Writing Experiences.” Roundtable participant at *CWPA*. 2013.
- “Graduate Student Perspectives on Early WPA Work.” Roundtable participant at *CWPA*. 2013.
- “Reframing Placement in Composition.” Research Network Forum participant at *CCCC*. 2013.
- “Haunted Places: Composing Possibilities for Democratic Design.” Panel chairperson at *CCCC*. 2013.
- “Reaching Out: Using Argument, Assessment, and Intra-Campus Collaboration to Make the Private Work of Composition Public.” Panel chairperson at *North Carolina Symposium on Teaching Writing*. 2013.
- “Bridging the Gap Between Placement, Instruction, and Our Values.” Panel participant at *North Carolina Symposium on Teaching Writing*. 2013.
- “Gateways to Dissent: Wisconsin Labor Protests, Civic Engagement, and Translingual Pedagogies.” Panel chairperson at *CCCC*. 2012.
- “Composition as a Site of Resistance to Cultural Assimilation.” Panel participant at *CCCC*. 2012.
- “Constructions of Disability in the Composition Classroom.” Panel participant at *CCCC*. 2011.
- “‘Allowing for Inclusion’: Addressing Military Students through Transformative Hospitality.” Presented at *University of Wisconsin-Milwaukee’s First Year Writing Symposium*. 2010.
- “I Read it on Your Wall: Facebook’s Influence on First-Year Writing Students’ Conception of Audience.” Presented at the *Popular Culture Association/American Culture Association National Conference*. 2010.
- “Online Peer Review as Reflective Practice.” Presented at *Marquette University Graduate Student Conference*. 2009.
- “Past as Present in Eugene O’Neill’s *Long Day’s Journey Into Night*.” Presented at *Saint Louis University’s Graduate Student Colloquium*. 2007.
- “‘They Perished in the Seamless Grass’: Emily Dickinson’s Indictment of the U.S. Civil War.” Presented at *Saint Louis University’s Graduate Student Colloquium*. 2006.

INVITED PRESENTATIONS

Forum on Access and Dis/Ability, English Department, UWM, 2013
 University TA Orientation Session, UWM, 2012, 2013
 Workshop on Language Conventions, English Department, UWM, 2012
 ESL Training Session, English Department, UWM, 2011
 Workshop on Editing, English Department, UWM, 2011
 Workshop on In-Class Discussions, English Department, UWM, 2010

ADMINISTRATIVE EXPERIENCE

Writing Program Assessment Committee, English Department, UWM, 2010-2013
 WPA-GO (Graduate Organization) Committee, elected member, CWPA, 2012-14
 English Graduate Policy Committee elected representative, UWM, 2012-13
 Writing Course Coordinator, ESL Program, UWM, 2013
 Composition Advisory Committee member, UWM, 2012-13
 English 095 (Fundamentals of Composition) Coordinator, UWM, 2012-13
 ESL Testing Alignment Committee, UWM, 2011-12
 Composition Curriculum Committee, UWM, 2011
 Graduate Faculty Committee elected representative, UWM, 2010-11

SERVICE

University of Wisconsin—Milwaukee

Pilot Assessment Project on Writing & Critical Thinking Committee, 2013-14
 Selection Committee, Common Reading Experience, 2013-14
 Planning Committee, Midwest Interdisciplinary Graduate Conference, 2013-14
 Discussion Leader, Common Reading Experience, 2013
 Teaching Award Selection Committee, English Department, 2013
 Rhetoric & Composition Mentor Match Program creator and co-chair, 2013
 University Writing Assessment Talk coordinator, 2012
 University Open House volunteer, UWM, 2011, 2012
 Virginia Burke Writing Contest Judge, English Department, 2011
 Rhetoric & Composition Elected Graduate Student Representative, 2010-11

Saint Louis University

International Graduate Student Conference coordinator, 2006, 2007
 Abanico Literary Magazine Editor, 2005-07

PREVIOUS EMPLOYMENT

Student Services Specialist, National-Louis University, 2008-09
 Prior Learning Assessment Specialist, AIU Online, 2007-08
 Communication Coordinator, Downers Grove Chamber of Commerce, 2003-05
 Childcare Provider; English and Math Tutor, 2002-03

LANGUAGES

Intermediate level of Spanish
 Proficient in HTML

PROFESSIONAL MEMBERSHIPS

College Composition and Communication
Council of Writing Program Administrators
National Council on Measurement in Education
National Council of Teachers of English

REFERENCES

Charles Schuster, Professor of English and Director of the Honors College, UWM
Mary Louise Buley-Meissner, Associate Professor of English, UWM
Patricia Mayes, Associate Professor of English, UWM
Lawrence Kuiper, Associate Professor of French and Interim Director of ESL
Programs, UWM
Peggy O'Neill, Professor of English and Director of Composition, Loyola
University Maryland
Diane Kelly-Riley, Assistant Professor of English and Director of Writing,
University of Idaho