

RETROSPECTIVE.

# THE POLITICS AND PUBLIC POLICY OF LARGE-SCALE WRITING ASSESSMENT

**Carolyn Calhoon-Dillahunt**

Yakima Valley College

For more than 150 years, standardized testing has been a part of the U.S. education system. Almost from the outset, standardized testing was inextricably linked to writing assessment and, thus, to writing instruction and, ultimately, to writing as a discipline. Early concerns about the “problem” of student writing revealed by standardized assessments resulted in increased attention to writing and writing instruction for teachers, for schools, and, eventually, for policymakers. As a result, for good and bad, writing (granted, often defined and assessed in reductive ways) holds a position of primacy in assessment and in educational policy, a position that garners attention and resources, but also scrutiny and intrusion.

In this section introduction, I briefly trace the history of large-scale writing assessment and how it has been entwined with politics and policymaking, situating the specific essays featured in Part Two of this collection in the “reform and accountability era” of large-scale standardized testing. From there, I discuss core themes around which these distinct articles coalesce: the policy intentions for and resulting uses and misuses of large-scale writing assessment in the 2000s; the consequences of mandated writing standards and high stakes writing assessments on curriculum, teachers and teaching, and students; and the possibilities enabled through some large-scale writing assessments.

## **A BRIEF HISTORY OF THE RELATIONSHIP BETWEEN LARGE-SCALE WRITING ASSESSMENT AND POLICY**

Although the purposes of standardized assessments have shifted over the past century and a half, gatekeeping and tracking have been primary among them. The earliest standardized tests focused on achievement of basic skills, such as language and literacy skills. Such tests were quickly taken up by selective colleges to determine admissions (National Education Association [NEA], 2020) and

placement into “remedial” writing coursework, starting with Ivy League schools in the late 1800s (Haswell, 2004). These early forays into writing assessment as gatekeeping planted the seeds of both basic writing and near universal first-year writing requirements in postsecondary study.

By the turn of the century, the founding of the College Entrance Examination Board meant that admissions testing and thus writing assessment became “outsourced,” and assessment became a professional, prolific, and profitable industry separate from the institutions that relied on their results (Huot, O’Neill, & Moore, 2010). In *Before Shaughnessey*, Ritter (2009) observes that accessibility of higher education, increasingly available to the masses after WWI and even more so with the GI Bill post-WWII, shifted the focus of writing assessment. Writing assessment became preoccupied with surface-level correctness, and remediation was prescribed to resolve students’ perceived lack of preparation for college-level writing. Over the course of the 20th century, writing placement also became increasingly disconnected from writing curriculum, as many institutions, especially open-admissions institutions, shifted from locally scored timed writing exams to externally scored standardized indirect writing assessments (Haswell, 2004).

In the 20th century, standardized testing expanded to assess proficiency, aptitude, intelligence, and more. However, according to Rosales and Walker (2021), “since their inception almost a century ago, the tests have been instruments of racism and a biased system,” founded on the pseudo-science, eugenics, and grounded in *white racial habitus* (Inoue, 2015). Nowhere is this racism more apparent than in standardized writing assessments. The purposes for such testing grew beyond simple gatekeeping for university admissions to diagnosing deficits, measuring skill sets, and predicting future performance. As a result, standardized testing was increasingly tied to educational decision-making (NEA, 2020), with the results of a single measure—generally an indirect measure embedded in White language and culture supremacy—being used to classify, rank, track, and exclude students. These approaches disproportionately affected historically underserved students, particularly students of color. Political support of large-scale testing as an important educational tool was sealed with the passage of the 1965 Elementary and Secondary Education Act. The first national assessment, the National Assessment of Academic Progress (NAEP), addressed in Applebee’s article in this section, was administered in 1969.

In the later 20th century, alarming reports of an impending literacy crisis, a crisis of “mediocracy,” and its implications for the U.S. economy, such as *Newsweek’s* “Why Johnny Can’t Write” (Sheils, 1975), *A Nation at Risk* (National Commission on Excellence in Education, 1983), and *Time for Results* (National Governors Association, 1985), led to calls for reform and accountability. These

calls for action resulted in a range of state-level policy solutions. One common action was increased implementation of statewide standards and assessment of students, from elementary to secondary-level, often in the form of direct assessments of writing and other basic skills. These standards and assessments were designed to impact curriculum and instruction and frequently were developed in response to employer demands, but with the influence of disciplinary experts. For instance, Sandra Murphy's (2003) *Journal of Writing Assessment* article, "That Was Then, This Is Now: The Impact of Changing Assessment Policies on Teachers and the Teaching of Writing in California," describes the California Assessment Program. This program developed in the early-1980s and was regarded as cutting edge for its focus on direct writing assessment. Murphy (2003) notes that half the states also were conducting direct writing assessments by the mid-1980s.

The essays in this section were published during a new era of large-scale assessment focused on educational "accountability." These approaches assumed test scores and high stakes could be used to raise standards. Literacy and writing remained key areas of concern and focus. By the late 1990s, many legislatures were moving toward holding schools and teachers accountable for improving students' performance on state-delineated standards, such as California's 1999 Public Schools Accountability Act (Murphy, 2003); however, the passage of the No Child Left Behind Act (NCLB) of 2001 mandated regular state-wide standardized testing coupled with financial performance-based penalties and rewards to push educational reform.

Although problems with one-dimensional accountability—and accountability resting entirely on the test scores of "hapless students" (White, 2005, p. 148)—were evident early on in K-12 education, this high stakes, testing-centered approach to educational accountability quickly "trickled up" to higher education. The 2006 Spellings Commission Report, which called for improving "accessibility, affordability, and accountability" in higher education, resulted in the 2008 Higher Education Opportunity Act, ushering in a wave of new accountability measures, increased federal regulation and data reporting requirements and a greater federal oversight role in institutional accreditation (Eaton, 2008).

Under the Obama administration, the accountability movement accelerated and increasingly gravitated toward the neoliberal economic policies of "paying for performance," what Toth, Sullivan, and Calhoun-Dillahunt (2016) describe as "a dubious method of improving educational outcomes through financial penalties and rewards already well-tested (and failing) in K-12 reform efforts" (p. 392). In elementary and secondary education, Race to the Top competitive grants, funded through the 2009 American Recovery and Reinvestment Act, helped propel states toward adopting the newly-minted Common Core

State Standards (CCSS), which Hammond and Garcia (2017) studied in their piece in this section. The English Language Arts and Mathematics Common Core, initiated by the Council of Chief State School Officers (CCSSO) and the National Governors Association (NGA) with the support of Achieve, Inc., were taken up by nearly every state (CCSS Initiative, 2022), often alongside the PARCC or Smarter Balanced online tests designed to measure these standards.

According to Adler-Kassner (2017), this accountability age has been driven by increased external influence on educational standards and outcomes by lawmakers, influential corporations, and many groups and actors that make up the reform-minded Educational Industrial Complex (EIC), who tell the story of “The Problem with American Education and How to Fix It” (p. 320). Toth et al. (2019) note, “over the last few decades, calls among both state and federal policymakers to improve student retention and degree completion have increasingly been framed as a matter of institutional ‘accountability’” (p. 2). According to Calhoon-Dillahunt (2018), the EIC’s solutions “privilege proficiency and efficiency (aka ‘success’ and ‘completion’) over learning and development” and their view of ‘accountability’ is market-oriented, with ‘value’ measured almost exclusively in economic terms” (p. 281). As a result, developmental and first-year writing are primary targets in “the EIC’s quest to streamline and economize higher education” (Calhoon-Dillahunt, 2018, p. 281). In the past decade, some states—Florida and Connecticut, for instance—have intruded into policies that were once institutionally determined, such as placement and developmental education, and most states have enacted performance-based funding policies in an attempt to drive reform.

## **ACCOUNTABILITY CONSEQUENCES AT STATE AND NATIONAL LEVELS**

The four chapters in this section are situated directly in the reform and accountability era. While the scale of “large-scale” and the policy implications—local, state, or national—vary with each assessment studied, the chapters together examine the intentions, politics, and misperceptions behind externally imposed writing standards and high stakes writing assessments and the resulting material and policy ramifications of these reform and accountability efforts.

In “The Misuse of Writing Assessment for Political Purposes,” Edward M. White (2005) identifies three focal areas of writing assessment that have been shaped by politics and public policy: high school proficiency testing, college placement, and mid-career assessments in colleges. The latter, “junior” writing assessments, which are addressed only in White’s piece, are comparable to high school proficiency testing in many ways. The remainder of the collection of

articles focus primarily on one of two significant and long-standing types of assessments White describes: secondary-level writing proficiency assessments and college writing placement testing.

In addition to White, Arthur N. Applebee and co-authors J. W. Hammond and Meredith Garcia all address K-12 writing proficiency testing and standards at the state and national level. Applebee's "Issues in Large Scale Writing Assessment: Perspectives from the National Assessment of Educational Progress," and Hammond's and Garcia's "The Micropolitics of Pathways: Teacher Education, Writing Assessment, and the Common Core" detail national writing standards and writing assessments and their consequences broadly. Applebee (2007) discusses the National Assessment of Educational Progress (NAEP), a congressionally mandated assessment across multiple subject areas, including a writing assessment, given to a representative sample of elementary and secondary students across the country. Applebee documents issues with large-scale writing assessments and the ways disciplinary expertise has been leveraged to improve the test and its utility. Hammond and Garcia (2017), on the other hand, focus on the Common Core State Standards (CCSS). Rather than analyzing the large-scale assessments associated with CCSS, PARCC, or Smarter Balanced (SBAC), they study how teachers navigate these common national standards in their own local contexts.

Along with White, co-authors Christie Toth, Jessica Nastal, Holly Hassel, and Joanne Giordano interrogate college writing placement in the age of high stakes. In "Introduction: Writing Placement, Assessment, and the Two-Year College," which is part of a *JWA* special issue on two-year college writing placement, Toth et al. (2019) outline how two-year college writing placement has become a particular target for educational reformers, which has resulted in a reconsideration of the role of placement and common placement practices.

Collectively, these four chapters coalesce around three core themes:

- The intentions behind and (mis)use of mandated writing standards and assessments for accountability purposes.
- The consequences of large-scale, high stakes writing assessments on curriculum, teachers, and students.
- Positive outcomes and spaces for possibility among some large-scale writing assessments and the policy implications.

## **INTENTIONS, ASSUMPTIONS, AND (MIS)USES**

Educational reforms and policies are often well-intended, but how they are enacted and enforced is often troubling and troublesome, especially in the

“accountability era.” In their articles in this section, the authors share that intentions behind common writing standards and standardized assessments often seem reasonable and even laudable. For instance, White (2005) asserts that it is entirely logical to expect high school students to demonstrate a certain level of reading and writing skill upon graduating. High school writing standards and accompanying writing proficiency tests are promoted as a way to prepare students for postsecondary writing. Hammond and Garcia (2017) describe how definitions of “preparedness” became codified in the Common Core State Standards, enabling measurement of this elusive idea of “college and career readiness.” According to the Common Core State Standards Initiative (2021) website, a consistent, nationwide set of standards can be used to articulate and measure student progress and to ensure students have acquired the necessary skills and knowledge to achieve success in postsecondary education and the workforce. Standardized testing, then, is viewed by policymakers and others involved in education reform as a way of raising standards and monitoring progress. According to the 2004 National Commission on NAEP report, a “high school diploma was no longer the culminating degree for most students” (Applebee, 2007, p. 86). Applebee also observes that about half of high school students who continued on to college were placed into developmental education, suggesting that many students were graduating from high school underprepared to do the sort of writing required in higher education. Thus, assessing 12<sup>th</sup> graders’ readiness for college, military, and career seems essential.

According to White (2005) and Toth et al. (2019), in some ways, placement testing aligns with intentions for high school writing proficiency testing, ensuring students are “ready” to do college work. The theory behind placement assessments is to match students to appropriate coursework, which allows college writing programs to maintain high standards in first year writing while providing support for underprepared students before or as part of their first-year writing coursework (White, 2005). In their article, Toth et al. (2019) share Willingham’s 1974 algorithm for understanding the role of placement assessments, a logic still pervasive in placement and developmental writing today. This logic suggests that, by identifying students with poor writing skills and matching those students to coursework designed to improve those skills, student learning and retention in writing courses will be improved.

Holding institutions accountable for student learning and achievement is also reasonable, according to White (2005): “it is wholly appropriate for politicians and citizens to inquire into whether the schools are accomplishing established goals” (p. 25). After all, states and local taxpayers, in particular, invest heavily in education, and they should expect students to graduate with the knowledge and skills needed for postsecondary pursuits. However, as Toth, Sullivan, and

Calhoun-Dillahunt (2016) have observed, accountability measures often fail to acknowledge that “the academic playing field is not level. An institution’s record of ‘success’ is largely shaped by its student demographics and resources” (p. 401). Moreover, high stakes measures offer limited information about student achievement and potential, yet are used, often singularly, to make consequential educational decisions.

In the “accountability era” of education, reforms are enforced through high stakes assessments. Problematically, accountability for these educational reforms is one-dimensional and one-directional, with consequences for schools (and thus students), regardless of their capacity and resources. White (2005) questions this one-way accountability that holds teachers and schools responsible for students’ performance on a single assessment without consideration of other influential factors, including school environment, quality and experience of teachers and administrators, learning support for students and teachers, among others, and, importantly, without consideration for policymakers’ own responsibility to ensure equal access to education and to appropriately support and fund basic education as well as their ambitious new educational initiatives.

Regardless of how well-intended, education reform in the “accountability era” is too often driven by oversimplified perceptions and a lack of understanding of what motivates, creates, and indicates change. Hammond and Garcia (2017) observe that education reform typically tries to “manage educational pathways,” using standards and assessments to regulate how students move through the educational system and “in the process, managing student advancement, opportunity, and attainment.” However, they note, “educational complexity is not so easily tamed,” and, ultimately, “[r]eform initiatives can only standardize so much” (Hammond & Garcia, 2017, p. 2). High stakes assessments enter the equation under the assumption that financial penalties and rewards will inspire desired reforms and create desired results. Linda Darling-Hammond (2007) asserts that accountability-oriented policies like NCLB misidentify the problems in education, assuming that “what schools need is more carrots and sticks rather than fundamental changes.” In an *NPR* interview, NCLB cheerleader turned outspoken critic Diane Ravitch adds that “measure and punish” is not an effective way to prompt change: “incentives and sanctions may be right for business organizations, where the bottom line—profit—is the highest priority, but they are not right for schools” and, in fact, have led to manipulation, dishonesty, and even cheating as schools compete for or try to preserve scarce resources (Inskip, 2010).

Not only is educational reform founded on misperceptions about how to implement change, but the writing assessments used to measure intended changes are based on fundamental misunderstandings about writing and how students

learn to write. Linda Adler-Kassner (2017) argues that “this lament, this story that students ‘can’t write,’ works from the premise that writing is ‘just writing.’ It’s a thing that writers bang out. It is constituted of words that are clear, that mean the same thing to everyone, that are easily accessible and need only to be plugged into forms” (p. 317). Toth et al. (2019) describe the foundational logic of traditional writing placement in much the same way; it’s built on the notion that such writing skills are attainable, measurable, and relevant to subsequent college-level writing coursework and that assessing these generic skills—and placing students accordingly—will lead to improved writing. In describing the development of the revised framework for the 2011 NAEP writing assessment, Applebee (2007) references a range of scholars who have challenged the “traditional emphasis on writing as a generic skill, taught primarily in English language arts or composition classes, and assessable through generic writing tasks detached from particular disciplinary or socially constituted contexts” (p. 163), yet the myths that “writing is just writing” and that “good writing” can be measured by a single test and without regard to context persist.

Raising the stakes on writing assessments and at the same time basing such assessments on fundamental misunderstandings about writing, assessment, and accountability has led to misuse rather than reform. For instance, the perception of writing as a generic skill has led to assessment tools that are often built to prioritize ease of measurement rather than achievement of higher order skills, resulting in assessments that focus on editing skills or formulaic writing tasks (Applebee, 2007). According to Toth et al. (2019), “The widespread reliance on commercially produced [writing placement] tests that measure a very limited construct of writing has prioritized knowledge of Edited American English conventions at the expense of any other outcome, primarily because these are the skills that can be easily measured” (p. 219). Thus, the tools that determine whether consequences will be meted out do not capture the lofty goals of the reform movement, and they are also biased against historically marginalized and minoritized students *by design*, essentially ensuring that the schools that serve such students will be penalized. These misuses are costly, in all senses.

In some cases, the high stakes assessments work against the very reforms they are trying to institute, case in point, high school writing proficiency testing. As several authors in this section articulate, the intentions behind large-scale high school writing assessments are to raise standards and increase student proficiency in writing for their postsecondary pursuits, as writing is a perceived “problem” despite the fact that high school graduation rates are over 85 percent and about two-thirds of those students enroll in postsecondary education after high school (U.S. Department of Education, 2021). To “inspire” students and teachers to take these standard-raising writing assessments seriously, many states tie earning



diplomas to passing state-mandated tests. Inevitably, implementing policies to solve one perceived problem, students' lack of preparedness for postsecondary pursuits, created many others. Policymakers were unprepared to admit that large portions of graduating seniors didn't demonstrate proficiency (White, 2005), although, given the frequency of testing students in K-12, they had fair warning about the likely results. Paradoxically, with the proliferation of dual enrollment programs in high school (NACEP, 2019), it's entirely possible for a student to simultaneously succeed in postsecondary coursework—and even earn a postsecondary degree—in high school, while simultaneously failing single-measure assessments designed to certify a student's "college-readiness."

As a result, grade 12 assessments are now given earlier in students' academic career, to allow more time for remediation and retakes. Assessments have been simplified to increase pass rates; instead of raising the bar, the assessments now represent the minimum level of competence required, and, even then, some students may not be able to pass them, so, according to White (2005), "exemptions, exceptions, and fraud enter the assessment system" (p. 146). Ultimately, these assessments create a Catch-22: students are deemed "unprepared" for postsecondary writing, although there is little consensus about what "college and career-ready" writing means (Applebee, 2007), by high school proficiency tests and writing placement tests, assessments largely disconnected from the writing curriculum. The number of "unprepared," as defined by student performance on these same high stakes assessments, leads policymakers to demand greater accountability, using high stakes assessments as the measure and mechanism for change.

## CONSEQUENCES

Attaching penalties and rewards to student performance on single assessment measures in order to drive educational reform and accountability policies has had far-reaching repercussions. The authors in this section address the negative consequences that have resulted from the use of mandated standards and high stakes writing assessments in three particular areas: curriculum, teachers and teaching, and students.

### IMPACT ON CURRICULUM

One of the most well-studied consequences of high stakes standardized testing is its impact on curriculum. Sandra Murphy (2003) notes high stakes assessments do not just measure achievement; they define it. Several authors in this section observed the ways that such assessments narrow, constrain, and distort writing

curriculum. Applebee (2007) argues that attempts to shape curriculum and assessment around abstract notions of “career and college readiness” have generally resulted in “a system of curriculum and assessment that focused on basic skills or on generic workplace tasks (e.g., business letter format) that easily degenerated into formulas with little real-world relevance” (p. 167).

The curricular impact of high stakes assessments can also be seen in postsecondary writing placement. According to Toth et al. (2019), “[i]n the nation’s open-admissions two-year colleges, where students enter from a wide range of academic trajectories and often have not taken any kind of admissions exam, placement assessment is nearly universal” (p. 215), and the use of commercial placement products predominates. One of the results of this sort of placement mechanism is that most two-year colleges offer multiple levels of pre-college writing courses, which may be similarly disconnected from first-year writing curriculum, focused instead on the “basic skills” developmental writers seemingly lack, and which sometimes prohibit students from accessing other college-level courses outside of English. On the other end of the spectrum, some colleges may exempt high performing students from the first-year writing requirement altogether, which suggests that first-year writing curriculum is not about introducing students to a discipline, but, instead, teaching generic “writing” skills.

Writing assessments that are disconnected from a college’s first-year writing curriculum provide limited utility for authentic placement, but they send powerful messages about how the institution views and values writing. Toth et al. (2019) recognize that writing placement “is not a neutral action” (p. 218); it communicates particular values and ideologies that affect how students, local high schools, and others perceive writing, and as a result, it can impact both high school curriculum and perceptions about the role of developmental and first-year writing on college campuses. Simultaneously, commercial placement tests also fail to communicate anything particular about a writing program, the theory that underlies its curriculum, and the practices it values; such assessment instead perpetuate the narrow conceptions of writing many students bring with them from high school and the commonly held notion that first-year writing is a course they need to “get out of the way.” Additionally, writing curriculum is impacted, negatively and positively, by current reform movements that seek to limit and accelerate developmental writing offerings (Toth et al., 2019).

## **IMPACT ON TEACHERS AND TEACHING**

Externally mandated standards and high stakes writing assessments also have a profound impact on teachers and the teaching of writing. Murphy (2003) argues standardized testing has deprofessionalized teachers, constraining their

opportunities for professional growth, undermining their autonomy and professional authority, devaluing their expertise, and blaming them for poor student performance on tests. White (2005) asserts that high stakes assessments are politically motivated and are used in disrespectful and manipulative ways toward teachers. He notes that many teachers are wary of large-scale assessments “because it almost inevitably narrows and often reduces what they do to simple numbers that will be used against their students and them” (p. 144).

Postsecondary writing instructors do not face the same types of blame and control as their secondary-level colleagues, but the reliance on placement tests, particularly at two-year colleges and other open-admissions institutions, have contributed to the notion that developmental writing and even first-year writing courses do not require professionalized writing teachers. The use of standardized placement tools that deem many—even a majority of students “unprepared” for college-level writing has led to a proliferation of basic writing courses. These courses are often viewed and even taught as “basic skills” courses, as courses designed to “re-teach” what students should have already learned in high school and, thus, not worthy of much investment. Toth et al. (2019) argue that the disconnect between theory and practice in writing placement assessment also detracts from the professional status of faculty who teach developmental and first-year writing.

## IMPACT ON STUDENTS

While the studies included in this set of articles don't address the impact of high stakes testing on students directly, the implications are clear: students bear the brunt of the consequences of standardized writing assessments. There is a long history of using writing assessments to gatekeep and rank students, and the consequences are even greater for students, especially historically underserved students, when assessments are tied to diplomas for college-level access. White (2005) argues that “Each of these assessments [high school proficiency exams, placement tests, mid-career writing assessments] represents a gate through which students must pass if they are to gain access to the privileges and enhanced salaries of college graduates, and so they carry a particular social weight along with their academic importance” (p. 145). The negative impacts of accountability policies and high stakes assessments previously described, from penalizing already under-resourced schools to narrowing the curriculum and reducing teacher agency and professionalization, also affect the quality of education students receive.

Toth et al. (2019) discuss most directly the impact standardized assessments have had on students in the context of placement. The authors cite Haswell's work

on the lack of predictability of writing placement tests; this lack of predictability of success has been corroborated by many others to reveal the ramifications for under-placement, which can extend student costs and time to degree, and for over-placement, which can cause failure, which is also costly, time-consuming, and can result in academic penalties. The consequences for misplacement disproportionately affect historically underserved student populations (Toth et al., 2019). Additionally, performance-based funding policies can penalize open-admissions institutions for student performance, which may incentivize those institutions to limit or refuse entry to students who, based on their placement scores, seem unlikely to succeed and, thus, threaten the college's funding. This disparately impacts minoritized and marginalized students (Toth et al., 2016).

## POTENTIAL AND POSSIBILITIES

This section makes clear that high stakes standardized writing assessments have often been detrimental to teaching and learning, to public perceptions about writers and writing, and to educational policy decision-making. However, the enterprise of large-scale writing assessment has not been without utility and even, at times, positive effects. Several of the articles in this chapter provide examples of well-designed standardized writing assessments that, when used as intended and without adding penalties and rewards that subvert their aims, serve a productive educational purpose and have contributed to our understanding of writing and writing assessment. Hammond and Garcia's study shows that teacher involvement in developing and mediating standards and assessments creates conditions for assessments to be used in ways that inform and improve curriculum and instruction, which are precisely the goals of these educational policy reforms.

Applebee's review of the framework for 2011 National Assessment of Educational Progress reveals that constructing and revising large-scale assessments, especially in consultation with teachers and disciplinary experts, enables writing assessment to reflect and shape research and scholarship in writing studies. Applebee (2007) reports on significant questions the National Assessment Governing Board (NAGB) considered about how to assess student writing in a valid, fair, and purposeful way as it revised its writing assessments for 2010 and beyond. According to Applebee (2007), in preparation for the revised 2011 NAEP writing assessment, the NAGB addressed questions about everything from the types of writing to be assessed, the prompts to use to generate writing, and the aspects of writing achievement to be measured to computer-mediated writing, test-taking accommodations, and time allotments. Such thoughtful consideration of assessment content and design leads to more informed and informative

assessments, especially when the stakes for such assessments remain relatively low for students, teachers, and institutions.

Because, as Applebee (2007) indicates, NAEP also served as a model for many state-developed assessments, NAEP's conscientiously designed and theoretically grounded assessment in writing had reverberating and likely positive effects on other large-scale writing assessments. Granted, the NAEP assessment, which appears to have been largely replaced at the high school-level by the CCSS-connected Smarter Balanced and PARCC assessments, still struggles with its intended goal of assessing student writing in ways that inform "preparedness for postsecondary endeavors," likely impossible to measure within a single, standardized assessment. However, the results have provided a fertile ground for study, on a large scale, which has enabled the field of writing studies to evolve.

Of course, mandated writing standards and large-scale writing assessments largely remain externally directed and developed. However, Hammond and Garcia (2017) remind us that policies have to be put into practice: "Standards . . . are never as autonomous or agentic as sometimes imagined; they are largely contingent on interpretation and implementation by the very actors they are intended to coordinate and perhaps constrain" (p. 184); indeed, they continue, "reforms put in place are seldom as stable and standardized as intended" (Hammond & Garcia, 2017, p. 186). The fact that policy is not determinative, is "not so easily tamed," means that policy requires support and buy-in to be fully enacted. Policy implementation is also negotiated and navigated within particular contexts: "Homogenizing educational projects like the CCSS are always alloyed with heterogeneous local perspectives, assumptions, and aims. While perhaps obscured by standardizing efforts, local differences are not erased by them." (Hammond and Garcia, 2017, p. 185). These mediated spaces are places of possibility, enabling the tools of policy implementation to be productively adapted and providing agency for those involved in their implementation.

In their study of student teachers, mentor teachers, and field instructors at three midwestern high schools, Hammond and Garcia (2017) observed that, while all teachers involved in their study utilized CCSS in some way in their curriculum development, they used and assessed the standards in different ways and for their own purposes, tied to their own local contexts. Study participants tended to curate and even "retrofit" the standards, rather than adopt them outright, which enabled the participants to select and prioritize the outcomes that fit their curriculum and goals and their students' needs as well as to use low stakes, classroom-based assessment practices to determine mastery. The study revealed that, instead of finding CCSS restrictive, the participating teachers tended to use the standards as a rhetorical tool, "as a medium for managing communication with stakeholders and—by extension—signaling professional participation in

the collective enterprise of American education” (p. 4). Some found the CCSS provided a common language for teachers, students, and parents to facilitate teaching and learning in the discipline, and others found using this “professional lingua franca” validated their work to external audiences, whether administrators, community members, or policymakers. Hammond and Garcia’s study reveals that, while policymakers may devise standards, teachers are the ones who enact them; the possibilities of educational reforms are tied to teacher buy-in and teacher agency to implement such reforms in context.

Further, their study suggests that teacher agency in determining and designing curriculum and assessment in context facilitates “professional accountability,” as described by Linda Darling-Hammond. According to Darling-Hammond (1989), “[p]rofessional accountability” requires that teachers are knowledgeable and engaged practitioners, who participate collectively in all aspects of teaching and learning, including assessment and local decision-making. Professional accountability has much more potential to drive positive and lasting change than the “carrot and stick” approaches associated with “accountability era” reforms. Hammond and Garcia’s (2017) work reveals that when teachers have agency in curricular decisions and when they are not threatened with punitive consequences, teachers often view imposed standards and large-scale assessments favorably.

## CONCLUSION

Education reform’s “accountability” turn has often been framed in terms of “value added,” with value defined—and “accountability” enforced—through neoliberal economic ideologies. Ravitch argues this competitive, market-based approach is wrong for public schools, which should function collaboratively and should share what works with others (Inskeep, 2010). In the “reform and accountability” era, large-scale writing assessments have often enabled these competitive and punitive policies. However, Rose (2012) asserts that “our philosophy of education—our guiding rationale for creating schools—has to include the intellectual, social, civic, moral, and aesthetic motives as well. If these further motives are not articulated, they fade from public policy, from institutional mission, from curriculum development” (p. 185). Because it’s connected to policy, mission, and curriculum—and, in fact, should emerge from these areas, writing assessment is foundational to how we articulate and ascertain “value” in education, and the future direction of writing assessment should consider “value-added” from the broader perspective Rose identifies.

To this end, the chapters in this section suggest a range of possibilities for future research. White’s, Applebee’s, and Hammond and Garcia’s work all

recognize the critical role of teachers in education reform and reveal the importance of teacher engagement with standards and assessments and of assessments emerging from and shaping curriculum. As teachers are enactors of reform policies, more attention should be directed toward understanding the impact of education policies and large-scale assessments on their practice and the role of professionalization and “professional accountability” plays in facilitating educational reform. Such research may reveal that investing in the changemakers, teachers, rather than investing in large-scale assessment tools may yield better results. Additionally, few studies talk to *students* about the ways in which they are experiencing “accountability” reforms, particularly how such policies and high stakes assessments affect their development and self-perceptions as writers and their conceptions of writing.

Toth, Nastal, Hassel, and Giordano’s work highlights the importance of assessing assessment tools. The work of researchers that questioned the validity, reliability, and predictability of commonly used commercial placement tests has resulted in many institutions abandoning such tests in favor of local alternatives or reducing the stakes by using such tests as one consideration, among others, for placement. These studies also led to revisions in commercial products themselves, often including a direct assessment of writing, albeit computer-scored. Not only is it important to assess validity and reliability in large-scale writing assessments, Toth et al. remind us of the importance of assessing the fairness of writing assessment tools and methodologies, especially in large-scale and high stakes assessments. Given that standardized writing assessments are rooted in White Language Supremacy and ableism, studying the consequences of writing assessments, in particular the disparate impacts of such assessments, can provide direction for how to redesign and even reimagine writing assessment tools that attend to local contexts and value diverse students. Toth et al. argue—and I agree—that two-year colleges are important spaces in which to conduct this research, as two-year colleges serve diverse students and communities and, with their open admissions policies, often serve as the primary access point for post-secondary education for the least advantaged students.

Finally, writing assessment research is one key way to change the public narrative around writing and to help policymakers develop informed solutions to the educational problems they are trying to solve. Writing researchers and scholars can contribute by asking different questions that counter the predominant failure-driven narrative. For instance, how can writing assessments provide evidence that student writing isn’t a “problem” and instead highlight the rich and rhetorically conscious ways students language and compose in classrooms with professionalized teachers developing curriculum appropriate to local contexts and students’ needs? How can large-scale writing assessments account for the

varied ways students demonstrate proficiency and success, for instance, in considering multiple measures instead of single assessments? How do lower stakes assessments provide more meaningful information and yield more positive results? How can large-scale writing assessments provide evidence of “college and career-readiness” by centering rhetorical dexterity and situated language practices instead of facility with Edited American English?

In addition to researching in ways that change the dominant discourse around writing, writing researchers and scholars can also practice their own rhetorical dexterity by sharing writing research in accessible ways with public audiences and policymakers. In other words, it is incumbent upon writing researchers to “[find] ways to communicate our expertise to those outside of our discipline and [seek] opportunities to participate in public conversations about literacy education” (Calhoon-Dillahunt, 2015). Future writing researchers can also take a page from two-year college teacher-scholar-activists who view engagement in educational policy as a professional responsibility, which requires “undertak[ing] the public work of defending educational access, teaching for democratic participation, and advocating for practices and policies grounded in disciplinary knowledges” (Toth, Sullivan, & Calhoon-Dillahunt, 2019).

## REFERENCES

- Adler-Kassner, L. (2017). Because writing is never just writing. *College Composition and Communication*, 69(2), 317-340.
- Applebee, A. (2007). Issues in large-scale writing assessment: Perspectives from the National Assessment of Educational Progress. *Journal of Writing Assessment*, 3(2), 81-98. <https://escholarship.org/uc/item/0zx1m9fg>
- Calhoon-Dillahunt, C. (2015). Finding our public voice. In D. Cambridge & P. Lambert Stock (Eds.). *Structural kindness: Essays on literacy education in honor of Kent D. Williamson* (pp. 163-169). National Council of Teachers of English.
- Calhoon-Dillahunt, C. (2018). Returning to our roots: Creating the conditions and capacity for change. *College Composition and Communication*, 70(2), 273-293.
- Common Core State Standard Initiative. (2022). *Frequently asked questions*. <http://www.corestandards.org/about-the-standards/frequently-asked-questions/>
- Darling-Hammond, L. (2007). *Evaluating 'No Child Left Behind'*. SCOPE: Stanford Center for Opportunity Policy in Education. <https://edpolicy.stanford.edu/library/blog/873>
- Eaton, J. S. (2008). The Higher Education Opportunity Act of 2008: What does it mean and what does it do? *Inside Accreditation*. <https://www.chea.org/higher-education-opportunity-act-2008-what-does-it-mean-and-what-does-it-do>
- Hammond, J. W. & Garcia, M. (2017). The micropolitics of pathways: Teacher education, writing assessment, and the Common Core. *Journal of Writing Assessment*, 10(1). <https://escholarship.org/uc/item/392847v0>



- Haswell, R. (2004). *Post-secondary entry writing placement: A brief synopsis of research*. CompPile. <http://www.comppile.org/profresources/writingplacementresearch.htm>
- Huot, B., O'Neill, P., & Moore, C. (2010). A usable past for writing assessment. *College English*, 73(5), 496-517.
- Inoue, A.B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing for a socially just future*. The WAC Clearinghouse; Parlor Press. <https://doi.org/10.37514/PER-B.2015.0698>
- Inskeep, S. (2010). *Former 'No Child Left Behind' advocate turns critic*. NPR. <https://www.npr.org/templates/story/story.php?storyId=124209100>
- Murphy, S. (2003). That was then, this is now: The impact of changing assessment policies on teachers and the teaching of writing in California. *Journal of Writing Assessment*, 1(1). <https://escholarship.org/uc/item/1fg1531r>
- NACEP. (2019). *Fast Facts*. National Alliance of Concurrent Enrollment Partnerships. [https://www.nacep.org/resource-center/?\\_sft\\_resourcetypes=fast-facts](https://www.nacep.org/resource-center/?_sft_resourcetypes=fast-facts)
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. [https://edreform.com/wp-content/uploads/2013/02/A\\_Nation\\_At\\_Risk\\_1983.pdf](https://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf)
- National Education Association. (2020). *History of standardized testing in the United States*. <https://www.nea.org/professional-excellence/student-engagement/tools-tips/history-standardized-testing-united-states>
- National Governors' Association. (1986). *Time for results: The governors' 1991 report on education* (ED279603). ERIC. <https://eric.ed.gov/?id=ED279603>
- Ritter, K. (2009). *Before Shaughnessy*. Southern Illinois University Press.
- Rosales, J. & Walker, T. (2021). *The racist beginnings of standardized testing*. NEA. <https://www.nea.org/advocating-for-change/new-from-nea/racist-beginnings-standardized-testing>
- Rose, M. (2012). *Back to school: Why everyone deserves a second chance at an education*. The New Press.
- Sheils, M. (1975, December 8). *Why Johnny can't write*. Newsweek. [https://www.lectorda.com/uploads/2/3/2/5/23256940/why\\_johnny\\_cant\\_write\\_\\_newsweek\\_1975\\_\\_1\\_.pdf](https://www.lectorda.com/uploads/2/3/2/5/23256940/why_johnny_cant_write__newsweek_1975__1_.pdf)
- Toth, C., Nastal, J., Hassel, H. & Giordano, J. (2019). Introduction: Writing assessment, placement, and the two-year college. *Journal of Writing Assessment*, 12(1). <https://escholarship.org/uc/item/8393560s>
- Toth, C., Sullivan, P. & Calhoun-Dillahunt, C. (2016). A dubious method of improving educational outcomes: Accountability and the two-year college. *Teaching English in the Two-Year College*, 43(4), 391-410.
- Toth, C., Sullivan, P. & Calhoun-Dillahunt, C. (2019). Two-year college teacher-scholar-activism: Reconstructing the disciplinary matrix of writing studies. *College Composition and Communication*, 71(1), 86-116.
- U.S. Department of Education. (2021). *Public high school graduation rates*. Institute of Education Sciences, National Center for Education Statistics. <https://nces.ed.gov/programs/coe/indicator/coi>
- White, E. M. (2005). The misuse of writing assessment for political purposes. *Journal of Writing Assessment*, 2(1). <https://escholarship.org/uc/item/9hg796d1>