

2

The Secret Cure for What Ails Large-Scale Writing Assessment: Teachers and their Students

Bob Broad

ILLINOIS STATE UNIVERSITY, USA

Abstract / Resumen / Resumo

Teachers and their students should play leading, authoritative roles in designing large-scale assessments and judging students' writing. They are in the best position to generate high-quality evaluations using affordable, sustainable, and educationally beneficial assessment methods. The success of this new assessment regime—developing assessments more rigorous, more valid, and fairer than what currently predominates—will depend on investment of time and money in building teacher-led professional education networks around writing assessment. Instead of distorting and corrupting the work of teachers and students as reductive standardized tests have done, large-scale writing assessment could provide a powerful educational opportunity and responsibility for these teachers and students. We already educate and reward teachers for guiding and assessing our students' rhetorical accomplishments; a teacher- and student-led portfolio assessment culture would re-invest assessment resources where they provide the best possible return: in the professionalization of writing teachers and the rhetorical development of students.

Profesores y estudiantes deberían encabezar el diseño e implementación de evaluaciones de escritura a gran escala. Dichos actores son los más aptos para crear mediciones de calidad usando métodos asequibles, sostenibles y educacionalmente ventajosos. El éxito de este nuevo régimen evaluativo —más riguroso, válido y justo que los actuales— requiere invertir decididamente tiempo y dinero en construir redes profesionales educativas lideradas por profesores, en torno a la evaluación de escritura. En lugar de distorsionar y corromper el trabajo de profesores y estudiantes, como lo han hecho las pruebas estandarizadas, las evaluaciones de escritura a gran escala proveerían una poderosa oportunidad y

responsabilidad educativa. Ya educamos y premiamos a los profesores por orientar los logros retóricos de nuestros estudiantes; una cultura evaluativa basada en portafolios, liderada por profesores y estudiantes permitiría reinvertir recursos de medición allí donde reportan mayores rendimientos: la profesionalización de profesores y el desarrollo retórico de los estudiantes.

Professores e estudantes deveriam encabeçar a elaboração e implementação de avaliações de escrita em grande escala. Ditos atores são os mais aptos para criarem medições de qualidade usando métodos acessíveis, sustentáveis e educacionalmente vantajosos. O sucesso deste novo regime avaliativo —mais rigoroso, válido e justo do que os atuais— requer investir decididamente tempo e dinheiro em construir redes profissionais educativas lideradas por professores, em torno da avaliação de escrita. Ao invés de distorcer e corromper o trabalho de professores e estudantes, como têm feito as provas padronizadas, as avaliações da escrita em grande escala proveriam uma poderosa oportunidade e responsabilidade educativa. Já educamos e premiamos os professores por orientarem as conquistas retóricas dos nossos estudantes; uma cultura avaliativa baseada em portfólios, liderada por professores e estudantes permitiria reinvestir recursos de medição ali onde reportam maiores rendimentos: a profissionalização de professores e o desenvolvimento retórico dos estudantes.

No. The teachers should do that work themselves. . . We want to do it ourselves (Gray, 2009, p. 20)

Assessments of written literacy should be designed and evaluated by well-informed current or future teachers of the students being assessed (CCCC, 2014)

Every literate society asks (or should ask) itself a set of important questions about how its youth learn to write:¹

- How (using what strategies and processes) and what (in what rhetorical situations or genres) do our students write?
- How and how well have we (teachers of writing) prepared our students for the rhetorical demands of work, of university studies, and of life in democratic society?

1 In developing and refining this chapter, I benefitted from helpful reviews and suggestions provided by David Slomp, Natalia Ávila Reyes, and an anonymous reviewer. Their comments helped me clarify and strengthen my discussion significantly. Any shortcomings of my work are solely my responsibility.

- What knowledge, skills, and dispositions around composition do our students offer the world?

These are urgent questions to which large-scale writing assessment (LSWA) provides large-scale answers. The question on which this chapter is focused is: Who should design these large-scale writing assessments and who should carry them out? In case you didn't notice yet, my answer is presented in the sub-title of this chapter: Teachers and their students should play leading, authoritative roles in designing assessments and judging students' writing. They are in the best position to generate high-quality answers to the questions above using affordable, sustainable, and educationally beneficial assessment methods. Closely tied to the "who" part of this assessment argument is a "what" component: In place of standardized writing tests (e.g., timed impromptu exams) that drastically diminish the ancient and robust construct of the rhetorical arts, LSWA needs to (re-)take the form of portfolio assessments. The success of this new assessment regime—developing assessments more rigorous, more valid, and fairer than what currently predominates—will depend on substantial investment of money and time in building networks of teacher professional education around writing assessment issues. Producers of standardized tests, meanwhile, should be restricted to advisory roles in this crucial educational project.

Such a teacher- and student-led writing assessment culture is not, however, the status quo. In the United States and in many other societies, policymakers and the general public have outsourced responsibility for the assessment of students' rhetorical learning to testing corporations such as Cambridge Assessment International Education, the Educational Testing Service (ETS), Pearson, and ACT (formerly American College Testing). In fact, large-scale writing assessment generates billions of dollars in profits² annually for testing corporations and other measurement organizations—Pearson alone reports profits of \$546 million for 2018. This is a winning scenario for the corporations and a terrible loss for education. For purposes both fiscal and educational, we should instead invest that money in developing localized, teacher-led assessment approaches like the award-winning systems pioneered in the 1990s in Kentucky as part of the Kentucky Education Reform Act (KERA) and in

2 Contrary to what many people assume, non-profit corporations such as ETS and ACT can and do make substantial profits. They are simply required to use their profits to cover expenses (including salaries) to maintain their tax-exempt status. Note that non-profits are awarded tax-exempt status on the understanding that they serve the public interest, not private interests.

Vermont.³ (Note, however, that the commercial testing industry recaptured the Kentucky and Vermont assessment systems in the late 2000s). The defining features of these two important U.S. historical exemplars of teacher- and student-led LSWA are:

1. With support and guidance from their teachers, students develop portfolios featuring a rich and diverse collection of rhetorical performances. (E.g., writing in school disciplines such as science, history, and art; literary analysis; political opinion; technical and professional writing; poetry, fiction, or non-fiction personal narrative).
2. The teachers who work directly with the students are the lead mentors and judges of their students' writing in this assessment system. Students engage deeply in self-evaluation of their work.
3. Students' collections of writing are judged in relation to a carefully developed set of writing standards (see examples below).

Some historical perspective on relationships between writing education and writing assessment will be helpful to this discussion, but I will only have room to gesture briefly toward useful resources by which readers can enrich their knowledge of histories of writing assessment. Norbert Elliot's *On a Scale: A Social History of Writing Assessment in America* (Elliot, 2005) is a highly useful history rendered by a researcher closely familiar with ETS. Less well-known but equally important is the same author's *Henry Chauncey: An American Life* (Elliot, 2014). In both books, Elliot carefully traces the complex and often fraught relationships between the standardized testing industry and the work of students and teachers in the schools; teachers are more beholden to principles and practices of literacy and democracy than to those of psychometrics and finances. Usefully, as I noted in a review (Broad, 2015) of *Henry Chauncey*, Elliot provides a detailed portrait of the ways in which Chauncey—founder of ETS—consciously chose to abandon an educational career in favor of the promise of launching a testing enterprise of significant economic, political, and cultural power.

Edward M. White's *Teaching and Assessing Writing* (White, 1994) demonstrates the author's monumental efforts to translate the ideology and methods of large-scale assessment (developed by Chauncey, ETS, ACT, and others during the 20th century) to serve the needs of university writing programs. While I have been critical on multiple occasions of the extent to which White

3 For in-depth background on the Vermont writing portfolio assessment system, see Hewitt, 1995. Regarding the Kentucky system, see Hillocks, 2002. For a history of holistic scoring in general, see Haswell and Elliot, 2019.

accepts and implements some of the methods of commercial standardized writing assessment, I also recognize the enormity of his struggle to develop assessment methods that would appear legitimate in the eyes of psychometricians while also staying true to the knowledge and values of teachers of writing such as White himself (Broad, 2012).

Tensions and conflicts between test-makers and teachers of writing (and their students) have run long and deep. For more than a century, teachers of writing have been sounding alarms about the detrimental effects of standardized testing on the teaching and learning of the rhetorical arts (Huot & O'Neill, 2009, p. 5). Consistently, however, testing organizations have brought to bear the power that Chauncey pursued and developed—economic, political, and cultural—to win lucrative contracts wresting authority for judgments of students' writing abilities from teachers' hands. Appeals to making assessment “scientific” and “objective” (see Aronowitz, 1988) are the chief persuasive tools by which this commercial enterprise has flourished. Now, however, there is bountiful reason to believe we should re-direct our investments in assessment away from corporations and toward teachers and their students.⁴

As successful as commercialized large-scale writing assessment has been politically, culturally, and economically, it has always struggled mightily to fulfill its most important responsibilities: to provide meaningful, useful, and trustworthy answers to questions about students' writing like those with which I opened this chapter. As educational professionals uniquely positioned to design and implement the best possible evaluations, teachers can and should be the leaders and authorities in large-scale writing assessment, and those teachers' students should also play a strong role. Newton (2017) supports this view when he argues in support of “purpose pluralism” in assessment design; Newton observes that, unlike most measurement specialists, teachers inevitably approach assessment with a pluralistic awareness of educational purposes. In a parallel argument, Flórez and Sammons (2013) demonstrate the ways in which “assessment for learning” as a concept and a practice brings together formative and summative assessment dynamics in support of teaching and learning.

4 As I work on final revisions to this chapter, the world—and especially the United States—is mired in a global COVID-19 pandemic. In this context, many U.S. colleges and universities have, at least temporarily, dropped their requirements for standardized college admissions tests. This scenario creates a new, possibly temporary, opportunity to make the case for localized, teacher-led, classroom-based, portfolio assessment systems like those for which I am advocating here.

Figure 2.1 sketches some of the most pressing conflicts and questions about large-scale writing assessment in language that I believe both psychometricians and educators will understand and agree with.

<p>Challenges of large-scale writing assessment—Collecting evidence to demonstrate:</p> <ul style="list-style-type: none">• Validity How sound are the <i>inferences</i> we make based on the assessment results? How positive are the educational <i>consequences</i> of our assessments?• Reliability Is evaluative diversity responsible and within appropriate limits?• Fairness What populations are advantaged or disadvantaged by our assessments? (Note the connection to the <i>consequences</i> element of validity above.)• Costs How sound are our investments of educational resources?

Figure 2.1. Paraphrased and adapted from Standards for Educational and Psychological Testing (American Educational Research Association [AERA] et al., 2014)

The first voice I wish to bring to discussion of these challenges comes from a unique perspective. Todd Farley spent fifteen years working in the standardized testing industry in the U.S. In 2009 he published *Making the Grades: My Misadventures in the Standardized Testing Industry* (Farley, 2009). The book is simultaneously extremely grim reading for teachers of writing and a hilarious exposé of this industry’s foibles. Here is one burning question and answer Farley offers near the end of his book:

What does it really mean to entrust decisions about this country’s students, teachers, and schools to the massive standardized testing industry? . . . It means ignoring the conclusions about student abilities of this country’s teachers—the people who instruct and nurture this country’s children every single day. (Farley, 2009, pp. 241-2)

Note that in this excerpt Farley speaks directly to *both* of the key questions about validity that haunt the standardized testing industry. How sound are the *inferences* made about students’ rhetorical abilities based on timed-impromptu test responses judged by under-paid, under-trained readers or by computers compared to the judgments made by teachers who work *daily* with these stu-

dents? Farley's book helps to illustrate that the question is laughable. The author reports multiple instances in which testing corporations for which he worked engaged in outright fraud and deception. Even in the absence of such willful wrongdoing, the material circumstances of the testing situation (e.g., restricted time and absence of resources for research, response, and revision) render any resulting inferences profoundly dubious in comparison with judgments made by teachers who work with students over extended periods of time on a diverse range of projects, helping them to develop their understandings of rhetorical situations, to conduct research and analyses, and to compose and revise the multi-media texts by which students try to win their arguments.

Likewise, Farley shines light on the second key element of assessment validity noted in Figure 2.1: the educational *consequences* of standardized testing of writing. The status quo of outsourcing assessment devalues the pedagogy and judgment of the teachers into whose hands we entrust our children and their education. It's not just that the systematic disregard of teachers' insights, analyses, and judgments built into the standardized testing industry is demoralizing to teachers. It's also that those disempowering movements are based on false claims, on a calumny against teachers: that teachers are "biased" (whether for and/or against) in relation to their students. Maja Wilson makes a similar point in her book *Re-Imagining Writing Assessment: From Scales to Stories*:

Educational policy ignores teachers' voices in the national educational discourse, treating [teachers] as problems to be solved, variables to be controlled, and villains to be held accountable through standards and testing. (Wilson, 2017, p. 135)

Teachers know their students abundantly well: their knowledge of their students' abilities and performances is strongly contextualized across varied projects and thus rich with invaluable data. While objectivist frameworks count this rich knowledge as a corrupting "bias," there are alternative frameworks of inquiry and value within which this special knowledge of teachers can and must be valued as a precious, irreplaceable resource rather than an epistemological and axiological liability (see Gallagher & Turley, 2012; Moss, 1994).

One of the most thoroughly researched portraits of large-scale writing assessment in the U.S. is George Hillocks's *The Testing Trap: How State Writing Assessments Control Learning* (Hillocks, 2002). As the book's title suggests, Hillocks is troubled by the corrupting relationships he found between standardized testing and teaching writing. Hillocks and his research team studied state-wide writing assessments in five of the fifty United States. In general, Hillocks's group became alarmed by the ways in which—and the extent to which—the products of testing corporations (and state institutions who

employ the same classical psychometric methods) undermined and distorted teachers' and experts' efforts to improve writing instruction. The problem was particularly severe in Hillocks's (and my) home state of Illinois, where (at the time of the study) the statewide writing test (the Illinois Goals Assessment Program) was designed as a timed-impromptu exam. Hillocks calls special attention to the duress and constraints teachers feel when large-scale tests contradict what they know about teaching writing.

I have quoted Mrs. Stafford at length to illustrate the process of indoctrination by the state. She feels powerless to resist the pressures from her own [school] administration and the state." (p. 132)

Hillocks's portrait of teachers' "powerlessness" to resist testing's "indoctrination" is vividly grim. However, LSWA need not be at odds with teachers' best teaching methods and professional judgments nor with our society's educational goals. To make large-scale writing assessment yield a positive return on our investments of education funds, we need to transform it into a system in which teachers and their students take on radically more—and more important—responsibilities.

A writing-assessment revolution like the one for which I am calling will not be easy to accomplish, and it will—based on the historical examples of Vermont and Kentucky mentioned above—be even more difficult to sustain in the face of powerful and sophisticated lobbying and marketing from testing corporations. As noted above, claims to objective and scientific assessment processes have been key marketing strategies for testing corporations since their inception. With the same gesture, commercialized test-makers have worked to cast aspersion on the value of teachers' knowledge about their students because that extensive knowledge is the single greatest threat to the commercial testing enterprise. After all, if we built our LSWA systems around teachers' rich and varied knowledge of their students' multi-dimensional rhetorical abilities, why would we waste money on a one-shot timed writing test?

The most useful piece of scholarship in considering relationships between these two frameworks (teaching vs. testing) is "Can There Be Validity without Reliability?" (Moss, 1994). As a parallel to Farley's book, Moss's innovative work in assessment theory comes from a scholar well versed and fully immersed in the discourses and frameworks of the testing industry. The author can see the strengths and weaknesses of these various frameworks, and makes the case for valuing each for its unique contributions. To paraphrase, Moss discusses two distinct and often opposing epistemological frameworks:

- Classical psychometrics (physical-science-based): Objectivist, con-

- text-free, unified, simplified, technical, statistical
- Hermeneutics (rhetoric-based): Interpretive, context-rich, diverse, complex, human judgments reached via collective critical conversation

The relevance of this analysis for the current discussion is obvious: Moss spotlights classical psychometrics as the predominant ideology guiding decisions about large-scale assessment, and then she carefully lays out the alternative hermeneutic tradition and discusses its potential to open up new possibilities for LSWA.

In a similar vein, Williamson (1994) notes three different approaches evident in the history of writing assessment. While the “factory model” has dominated since the rise of testing corporations in the early twentieth century, he highlights the educational assessment power of the even earlier “craft model.” Williamson explains how within the craft model, the long-term, context-rich, personal working relationships between teachers and students become a prized resource for answering our urgent questions about students’ learning and the quality of their rhetorical performances. The later success of the factory model of assessment depended in great part on de-valuing teachers’ knowledge of their students’ abilities.

Gallagher and Turley (2012) also champion teacher-driven writing assessment:

Can teachers take the assessment reins and lead into the twenty-first century, considering our already busy, sometimes overwhelming, professional lives? This book answers, resoundingly, *Yes, we can.* (p. 13; emphasis original)

Gallagher and Turley (along with Hillocks, Williamson, and Moss) firmly believe that with support from graduate education and collective practice in maintaining assessment standards (AERA et al., 2014), teachers can “take the reins” of large-scale writing assessment. They also believe that teachers *should* take this leading role for the sake of the quality of the answers only they can find to the crucial assessment questions with which this chapter began.

In this brave new world of teacher-led LSWA, do psychometricians and testing corporations become completely irrelevant? Not necessarily. Les Perelman (2018), for example, envisions limited, useful supporting roles for those who currently control, design, and carry out assessments: “The role of psychometricians should be limited to technical issues, and teachers and writers should constitute the final authority on issues of validity and reliability” (p. 40).

Given the historically lopsided power dynamics between educators and testing corporations, the role of psychometricians would need to be very care-

fully monitored. But I agree with Perelman that these assessment technicians could play a crucial role, bringing their distinct perspectives on key issues like validity and reliability. In other words, these powerful concepts need to be re-imagined along the lines that Lynne (2004) lays out in *Coming to Terms: Theorizing Writing Assessment in Composition Studies*. They need to be re-cast so that they are understood and practiced in terms that make sense for educators, students, and the general public, and then psychometricians can offer input from within their distinct framework (the one that currently predominates).

Teachers played the sort of leadership roles we've been discussing in the historical examples mentioned several times (Vermont and Kentucky), and they continue to put forward substantive and innovative proposals for new LSWA programs, even if their proposals have generally been ignored. For example, in 2004 a group of teachers (Boyd et al., 2004) researched and published an article in the statewide teachers' journal *Illinois English Bulletin* proposing to replace the "IGAP" timed-impromptu test that Hillocks critiqued with a statewide writing portfolio system (the Illinois State Portfolio Assessment Program, or "ISPAW"), in which teachers would undergo rigorous professional preparation and then evaluate their own students' work locally: "[We] propose a statewide portfolio assessment program as a more rigorous, fair, valid, and beneficial method for assessing writing in Illinois" (Boyd et al., 2004, p. 18).

In addition to working out the precise cost of judging each portfolio and comparing it favorably to the cost of the current timed-impromptu statewide test, these teachers made some interesting arguments in favor of their proposal. One argument was financial: "Tax money . . . will no longer go to out-of-state [testing] corporations" (p. 18). So, legislators and taxpayers who feel protective of their investments in education are reassured that dollars devoted to assessment are also dollars devoted to education, instead of the status quo where assessment is outsourced to the benefit of testing corporations.

Another argument for ISPAW is closely linked to the "keep your money local" line of thinking. Funds invested in teacher-led assessment are funds invested in teachers' professional education: "Illinois teachers will receive opportunities for meaningful and enriching professional development . . . through participation in the portfolio assessment system" (p. 20).

These arguments bring us to the heart of what makes teacher-led LSWA necessary. Public education is one of the key pillars of representative democracy, and most developed societies seem strongly committed to paying for a decent education for their children. The testing industry, by contrast, undergoes frequent paroxysms driven by the latest test and the latest set of standards. For example, over the past twenty-five years in Illinois, teachers have witnessed the arrival and departure of a cavalcade of tests used to gauge sec-

ondary students' writing abilities statewide: IGAP, ISAT, nothing, PARCC, SAT with Essay, and whatever comes next. I believe we are ready for a shift, in which assessment regains its stability and trustworthiness by re-integrating with the education system we are already funding. This is a call to stop the outsourcing of LSWA and instead to invest assessment funds in the professional educators already teaching our children and evaluating their writing. "The teachers should do that work themselves . . . We want to do it ourselves" (Gray, 2000, p. 20).

Testing corporations have served their own interests in part by persuading educational and legislative leaders that assessment is too complicated and too technical for teachers to handle. David Slomp, Board of Governors Teaching Chair at the University of Lethbridge in Canada and editor of the international journal *Assessing Writing* strikingly makes the opposite case: we need teachers to do this work precisely because it is so complex and *only teachers can understand* those complexities since they face them every day as they help their students learn and grow as writers:

The strongest argument for involving teachers in the process of large-scale writing assessment is the fact that when we measure writing ability, we are measuring something inherently unstable and complex—and when we take a developmental perspective, that complexity compounds exponentially. (Slomp, personal communication)

One of the most elegant and high-impact examples of teachers and students leading writing assessment is illustrated in Edward M. White's "The Scoring of Writing Portfolios: Phase 2" (2005). As most writing assessment scholars know, White devoted his career to helping colleges and universities design effective writing assessments, whether home-grown timed-impromptu tests or portfolio assessments like those implemented in the early 1990s at Miami University of Ohio, SUNY Stonybrook, and the University of Cincinnati. That was Phase 1. Phase 2 fits perfectly with the current discussion: Instead of sending students' texts or portfolios off to some other entity to be evaluated, in Phase 2 portfolio scoring, teachers and students work together intensively to investigate and document how and to what extent each student has (or has not) met key standards for writing performances.

There are several essential elements to Phase 2 portfolio scoring:

- Writing instruction proceeds with attention to whatever learning standards the institution deems most important, whether its own locally-developed set of standards (see Broad, 2003) or a profession-wide

set of standards such as the “WPA Outcomes Statement for First-Year Composition (3.0)” (Council of Writing Program Administrators, 2014) or the “Framework for Success in Post-Secondary Writing” (Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project, 2011).

- Writing assessment features teachers and students working together to develop portfolios in which a collection of student-authored texts demonstrates each student’s achievement of the designated standards.
- Students take on responsibility in their portfolio prefaces for carefully documenting and illustrating in detail how the components of their portfolios demonstrate their accomplishment of specific standards.
- The role of teachers then becomes a confirmatory one: Weighing in on how successfully each student-author documented and demonstrated their achievement of the standards.

White (2005) puts it this way:

Phase 2 scoring . . . reinforces the entire point of portfolios by making the assessor of first resort the student submitting the portfolio, who, in the reflective letter, performs the self-assessment that is the true goal of all academic assessment. The faculty assessment then focuses on that student assessment. (p. 594)

This approach vividly illustrates students working with teachers to meet standards set by their local school, district, or state or by the wider profession of writing studies. Notice that the process is rhetorically and meta-cognitively sophisticated and complex (documenting how one’s diverse writing performances show one’s mastery of learning standards) but not technically obscure in ways that would require psychometric expertise. The crucial research questions are straightforward and were presented in the opening paragraph of this chapter: “How and how well have our students learned to write?” Valid and reliable answers to those questions can take forms as simple as designations of “proficient” vs. “not-yet-proficient,” and the entire process of teaching, learning, and assessment can at last be coherent and mutually supportive.

Carrying forward the spirit of White’s vision of student-led assessment, Navarro, Ávila Reyes, and Gómez Vera (2019) designed a remarkable writing test for university admissions purposes in Chile. While I stand in admiration of multiple characteristics of their test—and the multi-layered theoretical model they developed in support of the test’s design—the principle most relevant to our discussion of teacher- and student-led LSWA is the authors’ insistence on giving students more freedom and responsibility than is typical

for standardized writing assessments:

standardized writing tests should place the student in a position of authority, autonomy, and agency over the subject matter addressed, where he or she can draw on his or her previous knowledge, skills, and experience to solve the task (Camp, 2009) in a design that promotes “self-authorship” (Broad, 2003). (Navarro et al., 2019, p. 8)

The approach I am advocating in this chapter is strongly allied with the principles and practices of these assessment researchers; my argument is to extend the high value they place on honoring students’ authority, autonomy, and agency to include teachers. I ask readers to envision what LSWA looks like when it acknowledges the authority, autonomy, and agency of students and their teachers.

I have presented a robust chorus of voices calling in unison for large-scale writing assessment to be brought under the control of classroom teachers and their students. If I have succeeded in persuading readers of this thesis, then another important issue needs to be addressed: the ways in which putting teachers in charge is a necessary but not sufficient condition for valid large-scale writing assessment. Two additional conditions must be met for the system to work optimally:

1. teachers must receive significant advanced education in teaching and assessing writing, and
2. the system must provide for what Pamela Moss calls “a critical dialogue” (p. 9) among professional peers as a warrant for evaluative decisions.

For starters, as Hillocks implored many years ago, teachers of writing need to undertake graduate study in the teaching of writing (writing studies, composition and rhetoric, etc.). Without significant knowledge of this field, teachers lack the necessary intellectual grounding and leverage from which to advocate for optimal teaching and assessment practices. Since in this chapter we are discussing how teachers can, should, and must take control of LSWA, we must add to Hillocks’s prescription graduate study in the sub-field of writing *assessment*.⁵

As discussed earlier regarding White’s “The Scoring of Writing Portfolios: Phase 2,” another crucial component of a teacher- and student-led writing assessment scheme is a wholesome and organically grown statement of what ed-

⁵ For example, my university offers a “post-baccalaureate certificate in the teaching of writing,” a series of six graduate courses that includes a course in writing assessment.

ucators believe students need to learn. On the contemporary scene, such statements are typically described as “learning outcomes” or “learning standards.” (Examples: CWPA Outcomes Statement; Framework for Success in Post-Secondary Writing Assessment; Common Core State Standards Initiative).

To clarify the point(s) at the heart of this discussion: When teachers pursue advanced education in writing studies and writing assessment and pair that education with meaningful and legitimate learning standards, then those teachers can guide their students through the process of learning how to succeed in a variety of rhetorical situations and learning how to document their successes (and shortcomings) as writers in relation to agreed-upon educational standards. Such a process provides everyone concerned—taxpayers, legislators, parents, students, teachers, and administrators—with robust, high-quality, trustworthy answers to our most urgent questions about students’ rhetorical development.

Once this classroom-based work has been completed, a larger assessment system must be engaged. For reasons similar to those noted by Slomp (personal communication) and Moss (among others), the perspective of the classroom teacher who has worked with a particular student must hold a “privileged” place in the assessment conversation by which that student’s work is judged. However, it is *not* satisfactory for that teacher to render judgment *alone*.

This brings us back to some of the claims (considered earlier) by which testing corporations persuaded educational systems to take assessment decisions out of the hands of teachers in the first place. Recall that the psychometricians promised to bring a scientific objectivity to testing processes. We now know that such claims are spurious at least in the context of literacy learning and rhetorical performances. The complexity and context-sensitivity of these activities generates dynamics by which interpretations and judgments will necessarily—and appropriately—diverge. Furthermore, principled differences among responsible judgments should be treated as a precious educational resource. “Objective assessment” is, in short, an oxymoron.

However, just because the offer of objectivity turns out to be illusory does not mean that teachers can or should be left to make high-stakes assessment decisions in isolation. To the contrary, Moss’s discussion of widespread hermeneutic assessment practices makes clear that hermeneutics offers and requires a process for warranting assessment decisions (for demonstrating, that is, the validity and fairness of those decisions) that is far more effective than standardized testing’s make-believe objectivity.

a more hermeneutic approach to assessment would warrant interpretations in a critical dialogue among readers that chal-

lenged initial interpretations while privileging interpretations from readers most knowledgeable about the context of assessment. Initial disagreement among readers would not invalidate the assessment; rather it would provide an impetus for dialogue, debate, and enriched understanding informed by multiple perspectives as interpretations are refined and as decisions or actions are justified. “The community of inquirers must be a critical community, where dissent and reasoned disputation (and sustained efforts to overthrow even the most favored viewpoints) are welcomed as being central to the process of inquiry” (Phillips). The point is to discriminate between blind and enabling prejudices by critically testing them in the course of inquiry. (Moss, 1994, p. 9)

Again, in fairness to a half-century of testing corporation marketing campaigns, they raised a legitimate concern: that the teacher “most knowledgeable about the context of assessment” might also be rendering judgments affected by “blind prejudices” whether for or against a particular student or group of students. Fortunately, hermeneutic assessment offers a methodology for sorting out the blind from the enabling prejudices: structured, critical conversations among knowledgeable professional peers (Broad, 1997; Moss, 1994).

Currently in the United States, the blind prejudice par excellence is racism. Many societies, including the U.S., are currently struggling openly with histories of racial violence and oppression rooted in slavery, and seeking out possibilities and strategies for achieving racial justice. Within the sub-field of writing assessment, substantial and important work is ongoing on this topic (Inoue, 2015; Inoue & Poe, 2012; Poe et al., 2018). As with all other questions of justice and validity, teachers will need to educate themselves about racism in writing assessment and work together to struggle against that history and avoid future instances. I offer this bracing example as a reminder to readers that addressing the shortcomings of current large-scale writing assessment requires that teachers and students play leading roles, but putting those teachers and students in charge will mark the beginning, not the end, of the project of creating high-quality assessment programs to support rhetorical education.

Recall that in *The Testing Trap*, Hillocks lamented the “indoctrination” of teachers and the distortion of the teaching of writing that resulted from statewide timed-impromptu tests like the IGAP test in Illinois. A teacher- and student-led writing assessment culture (at the state or the national level) would turn such distortion and indoctrination on its head. Learning what is required of them to lead such assessment systems would be a highly valuable

educational enterprise for both teachers and students. Teachers will learn histories, theories, and practices of teaching and assessing writing; students will grapple seriously with statements of rhetorical learning outcomes and work to analyze their own writing in relation to those outcomes.

In response to a draft of this chapter, a reader offered this pocket-sized sketch of the key steps in the proposed transformation of our cultures of writing assessment:

1. Help the public and political stakeholders understand the negative consequences of the status quo: undermining good teaching, confusing students about what makes writing successful.
2. Educate all stakeholders about the advantages of a teacher- and student-led model of writing assessment, based on theories and practices drawn from writing studies and rhetorical traditions.
3. Promote teachers' knowledge and expertise regarding the writing construct, writing pedagogies, and assessment theories. Invest money, time, and education in building teachers' expertise.
4. Build on the model of the National Writing Project, in which "teachers teach teachers" and support, critique, and enhance each other's teacher-research.
5. Continually gather and analyze systematic evidence of the various consequences (educational, ethical, political, economic, etc.) of this new assessment culture.
6. Innovate new solutions as they become needed and available.

The assessment revolution for which the voices highlighted here are calling will restore coherence to an educational ecosystem that has been disrupted for at least half a century: Instead of distorting and corrupting the work of teachers and students as reductive standardized writing tests have done for decades, large-scale writing assessment will once again provide a powerful educational opportunity and responsibility for those teachers and students. We already educate and reward teachers for guiding and assessing our students' diverse rhetorical accomplishments; a teacher- and student-led portfolio assessment culture will re-invest assessment resources where they provide the best possible return on that investment: in the professional education of writing teachers and the rhetorical development of our students.

References

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educa-*

- tional and psychological testing*. American Educational Research Association.
- Aronowitz, S. (1988). *Science as power*. University of Minnesota Press.
- Boyd, M., Broad, B., Conant, B., Freaner, M., Kixmiller, L., Nussbaum, A., Parlier, S., Spangler, S.B., & Wike, C. (2004). Making statewide writing assessment rigorous, valid, and fair: The Illinois State Portfolio Assessment of Writing (IS-PAW). *Illinois English Bulletin*, 91(3), 11-28.
- Broad, B. (1997). Reciprocal authorities in communal writing assessment: Constructing textual value within a "new politics of inquiry." *Assessing Writing*, 4(2), 133-167. [https://doi.org/10.1016/S1075-2935\(97\)80010-4](https://doi.org/10.1016/S1075-2935(97)80010-4)
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Utah State University Press.
- Broad, B. (2012). Mapping a dialectic with Edward M. White (in four scenes). In N. Elliot & L. Perelman (Eds.), *Writing assessment in the 21st century: Essays in honor of Edward M. White* (pp. 259-269). Hampton.
- Broad, B. (2015). Book review: *Henry Chauncey: An American life* by Norbert Elliot. *Journal of Writing Assessment*, 8(1).
- Conference on College Composition and Communication. (2009). *Writing assessment: A position statement*. National Council of Teachers of English.
- English Language Arts Standards: Writing. (n.d.). *Common Core State Standards Initiative*.
- Council of Writing Program Administrators. (2014). WPA outcomes statement for first-year composition (3.0). *WPA: Writing Program Administration*, 38(1), 129-143.
- Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project. (2011). *Framework for success in postsecondary writing*.
- Elliot, N. (2005). *On a scale: A social history of writing assessment in America*. Peter Lang.
- Elliot, N. (2014). *Henry Chauncey: An American life* (1st new ed.). Peter Lang.
- Farley, T. (2009). *Making the grades: My misadventures in the standardized testing industry*. Berrett-Koehler Publishers.
- Flórez, M. T., & Sammons, P. (2013). *Assessment for learning: Effects and impact*. University of Oxford.
- Gallagher, C. W., & Turley, E. D. (2012). *Our better judgment: Teacher leadership for writing assessment*. National Council of Teachers of English.
- Gray, J. (2000). *Teachers at the center: A memoir of the early years of the National Writing Project*. National Writing Project.
- Haswell, R., & Elliot, N. (2019). *Early holistic scoring of writing: A theory, a history, a reflection*. Utah State University Press.
- Hewitt, G. (1995). *A Portfolio Primer: Teaching, Collecting, and Assessing Student Writing*. Heinemann.
- Hillocks, G. (2002). *The testing trap: How State writing assessments control learning*. Teachers College Press.
- Huot, B., & O'Neill, P. (2009). Introduction. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 1-9). Bedford/St. Martin's.

- Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. The WAC Clearinghouse; Parlor Press. <https://doi.org/10.37514/PER-B.2015.0698>
- Inoue, A. B., & Poe, M. (2012). *Race and writing assessment*. Peter Lang.
- Lynne, P. (2004). *Coming to terms: Theorizing writing assessment in composition studies*. Utah State University Press.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5-12. <https://doi.org/10.2307/1176218>
- Navarro, F., Ávila Reyes, N., & Gómez Vera, G. (2019). Validez y justicia: Hacia una evaluación significativa en pruebas estandarizadas de escritura. *Meta: Avaliação*, 11(31), 1-35.
- Newton, P. E. (2017). There is more to educational measurement than measuring: The importance of embracing purpose pluralism. *Educational Measurement: Issues and Practice*, 36(2), 5-15. <https://doi.org/10.1111/emip.12146>
- Perelman, L. (2018). *Towards a new NAPLAN: Testing to the teaching*. NSW Teachers Federation.
- Poe, M., Inoue, A. B., & Elliot, N. (Eds.) (2018). *Writing assessment, social justice, and the advancement of opportunity*. The WAC Clearinghouse; University Press of Colorado. <https://doi.org/10.37514/PER-B.2018.0155>
- White, E. M. (1994). *Teaching and assessing writing* (2nd ed.). Jossey-Bass.
- White, E. M. (2005). The scoring of writing portfolios: Phase 2. *College Composition and Communication*, 56(4), 581-600.
- Williamson, M. (1994). The worship of efficiency: Untangling theoretical and practical considerations in writing assessment. *Assessing Writing*, 1(2), 147-173. [https://doi.org/10.1016/1075-2935\(95\)90021-7](https://doi.org/10.1016/1075-2935(95)90021-7)
- Wilson, M. (2017). *Reimagining writing assessment: From scales to stories*. Heinemann.