CHAPTER 5

TEACHING WRITING MATTERS

At its core, this book is about the competition to govern, measure, and exploit literacy as it has played out since 2006, the year we are using as our pivot point as we look backward and forward in order to make decisions about shaping the path ahead of us through local and national efforts. Perhaps most importantly, it is about asserting the primary role of teachers as powerful *sponsors of literacy* working through networks on numerous levels—a role that requires a renewed commitment to writing instruction and research in our kindergarten through college classrooms horizontally across the curriculum and vertically through grade levels.

In *Literacy in American Lives,* Deborah Brandt defines sponsors of literacy as:

> any agents, local or distant, concrete or abstract, who enable, support, teach, and model, as well as recruit, regulate, suppress, or withhold, literacy—and gain advantage by it in some way. . . . Sponsors are a tangible reminder that literacy learning throughout history has always required permission, sanction, assistance, coercion, or, at a minimum, contact with existing trade routes. (19)

While it is clear that more stakeholders than ever can be counted among the sponsors of literacy, so, too, is it clear that teachers are still among the primary sponsors of literacy. Furthermore, it may be that the only way for teachers to be effective sponsors of literacy at this current moment is if we do a better job of finding ways to shape and control the trade routes over which literacy travels. By trade routes, we mean those networks of pathways and stoppages through which literacy does, or doesn't travel. Brandt's work becomes vitally important in this effort as she helps us understand how literacy trade routes have been largely seized and governed by private economic interests:

> Literacy is being sponsored in much different ways than it was in the past. Through most of its history, literacy was affiliated with a few strong cultural agents—education, religion, local commerce. It tended to be learned in the same contexts in which it was intended to be practiced. Now, sponsors of literacy are more prolific, diffused, and heterogeneous. . . . Commercial sponsors abound. (197)

For example, as Brandt demonstrates, in the early days of America, the church was one trade route to literacy, as Sunday School was begun to teach poor children math and reading in addition to religious values. We would argue today that testing companies and deep-pocketed private foundations control the trade routes to an unprecedented degree, acting as both a conduit to (at its best) and blockade to (at its worst) literacy acquisition.

Looking back at 2006, we now realize we were caught up in the midst of a newly energized clash of sponsors: "These clashes typically are between long-standing residual forms of sponsorship [e.g., university writing programs writ large] and the new, between the lingering influence of literacy's conservative history and its pressure for change" (Brandt 193). Indeed, our book is a call to teachers at all levels to do what they can to shape existing and emerging trade routes in ways that maintain the importance of writing as a public good, not a private interest, in the service of educational equity and opportunity. Doing so requires at least five significant changes:

1. Providing time in the work day for teachers to engage with local and national networks
2. Allowing significant amounts of time during the work day for collaborative planning and problem solving in one's department or school with other teachers, and sharing this work at the state and national level not only through yearly conferences, but sustainable structures such as the National Writing Project
3. Renewing our commitment to principles and practices of shared governance
4. Continuing insistence that open access to research data and results be required of private companies and public educational organizations so that more stakeholders have a voice in assessing the results of a research project and a voice in actions that might follow
5. Demanding that our school districts, as well as state and federal governments, not cede their historical role in providing for the means and direction of our public education system to private testing companies and advocacy philanthropists

As we take stock of what we have learned from our research, the research of others, and emerging developments in writing studies, we are convinced more than ever that *the teaching of writing matters*. And, it matters that teachers of writing be involved in creating sustainable structures for change in the ongoing battle over literacy.

EMPOWERING SPONSORS

In 2011, asserting its traditional role as a sponsor, the U.S. Department of Education appointed twenty-eight education advocates, civil rights leaders, scholars, lawyers, and corporate leaders to its Equity and Excellence Commission. Their report was issued early in 2013. Like many who follow these issues, we expected this report to be more of the same—more competition, more charter schools, more testing, more privatization, more corporate sponsorship. Unexpectedly, the commission's report, *For Each and Every Child: A Strategy for Education Equity and Excellence*, advocates funding schools justly and equitably, providing well-qualified teachers in all schools, opening access to universal early childhood education, serving and supporting at-risk students and families in high poverty areas (including providing access to health care), and governing to promote excellence (The Equity and Excellence Commission). The Commission's five-point action strategy touches upon what we believe to be the opportunity that holds the most promise for empowering and engaging teachers as sponsors of literacy—meaningful professional development—although our version of professional development places teachers in a more active, authorial role than does that of the commission.

The report's introduction, excerpted below, echoes the same findings many of us have been struggling with for many years, and is one of the most powerful admissions that despite all of our reforms and all of our tests, little progress has occurred:

> In 1983, *A Nation at Risk* famously spoke of the "rising tide of mediocrity" that threatened our schools. Nearly 30 years later, the tide has come in—and we're drowning. Since that landmark report, we've had five "education presidents" and dozens of "education governors" who have championed higher standards, innovative schools, better teaching, rigorous curricula, tougher testing and other education reforms. And, to be sure, there has been important progress. Reading and math performance levels in our elementary schools, for example, have improved in recent years, as has mathematics performance in our middle schools. (14)

Note that, once again, writing does not even merit mention in this account. The commission goes on to state:

> Except in a few states, however, the incremental steps we have taken have not been enough to keep pace with the dramatic

> improvement other nations have made in their school systems. Moreover, any honest assessment must acknowledge that our efforts to date to confront the vast gaps in educational outcomes separating different groups of young Americans have yet to include a serious and sustained commitment to ending the appalling inequities—in school funding, in early education, in teacher quality, in resources for teachers and students and in governance—that contribute so mightily to these gaps. (14)

For the remainder of this chapter we'd like to focus on one very specific aspect of this commission's action plan—teacher quality—within the specific concerns of literacy. For us, the question is not simply, how do we improve teacher quality at this moment in time. Rather, as we frame the question—because the teaching of writing matters, and teachers of writing matter—we must ask how teachers can now assert a primary role as sponsors of literacy who are enabled to shape the trade routes along which literacy travels. For us, the answer lies not in the hiring of more, better teachers for more, better money. Instead, the answer lies in context-specific professional development work that empowers and engages teachers as sponsors of literacy in both local and national networks. And, in spite of its critics (which sometimes includes us), we do believe that the new Common Core State Standards can be a motivating force in this effort. As the commission notes: "the recent formulation of Common Core State Standards . . . provides a unique moment to leverage excellence and equity for all and to build on efforts to foster critical thinking and problem-solving, creativity and innovation, and communication" (15). However, this will *only* be true if teachers, beginning in kindergarten and through college, are empowered and engaged in shaping the routes CCSS travels, instead of allowing corporate interests to reign as cash-strapped states are lured into adopting questionable practices.

One such practice is the continued overreliance on standardized tests as the basis of school reform. As we evidenced earlier, despite hundreds of years of testing that has little to show in the way of improved learning, we persist in using them as a primary measure and motivator. Recently, this practice has shown great potential to be even more detrimental to reform efforts due to the growing insistence on machine-graded scoring of writing. Current machine-graded scoring of standardized tests of writing is purported to be aligned with CCSS, but in fact is rooted in simplified prompts and short answer essays that can be measured by machines on the most basic levels. Grand claims about the cost savings and reliability of machine-graded scoring, as well as the willingness of cash-strapped states to adopt these programs, are not new. Nor is the repeated

abandonment of these testing products when they fail in exactly the ways that literacy scholars predict they will fail. For example, in 2002, Indiana adopted ETS's E-rater scoring engine. They quickly realized its shortcomings:

> It couldn't reliably handle questions that required students to demonstrate knowledge from the curriculum. State testing officials tried making lists of keywords the software could scan for: in history, for example, "Queen Isabella," "Columbus," and "1492." But the program didn't understand the relationship between those items, and so would have given full credit to a sentence like, "Queen Isabella sailed 1,492 ships to Columbus, Ohio." Cost and time savings never materialized, because most tests also had to be looked at by human graders. (Goldstein)

In a recent and widely publicized study of the accuracy of machine-graded scoring of human writing, Mark Shermis (University of Akron) and Ben Hamner (Kaggle) compare the abilities of nine different scoring engines to rate student writing. These authors found that "overall, automated essay scoring was capable of producing scores similar to human scores for extended-response writing items with equal performance for both source-based and traditional writing genre" (2). But a close look at their study reveals significant problems not only with their analysis, but also with what they analyzed. In his critique of this study, Les Perelman identifies four main areas of concern, which we summarize here:

1. The use of differing scoring rules for human graders and machine graders, which brings into question the validity of results. The claims made by Shermis and Hamner are based on the resolved score (RS). Many of us are familiar with resolved scores. For example, in writing programs with an exit portfolio, two readers will score a portfolio, and if their scores are identical or adjacent (e.g., do not differ by more than 1 point on a 6 point scale), then the resolved score is determined by adding the two scores and dividing them by two. If the two scores differ by more than 1 point then a third reader is used to determine the RS. Shermis and Hamner, however, not only use scoring rules for human graders that are not in line with best practices, but also use different scoring rules for the machine graders, thus using two different variables in their comparison: Shermis and Hamner's "text uses the variable H1H2, the reliability between the two readers, as the measure for reader reliability, while the measure for machine performance is reliability between machine and the resolved score (RS)" (5). Perelman considers this to be the greatest problem with their study.

2. The lack of standardly expected statistical tests appropriate for the data. Without using commonly expected statistical tests, some results seem to be based more on "hunches" or inferences rather than on valid statistical measures of significance.
3. The failure to test the entire model for significance. Without doing so, there is little way to prove that machines didn't outperform humans simply by random chance or pure dumb luck.
4. The lack of consistency in what was actually being measured. Half of the data sets were not extended written response essays, but rather were one-paragraph responses. Further, four of the datasets were not designed to measure writing ability, but rather reading comprehension and literary analysis. The difference in the length of the samples and the fact that many were not designed to measure writing ability did not stop Shermis and Hamner from using them to make claims about the accuracy of machine graders to score writing ability.

Finally, Perelman concludes, not only do Shermis and Hamner fail to prove their conclusion, but rather, "the data support the assertion that human scorers performed more reliably than the machines on the longer traditional writing assignments" (3).

While there is much we find troubling about this research, what troubles us most about Shermis and Hamner's study in light of our work here is Perelman's last critique—much of the writing being analyzed involved one-paragraph responses—and many were not even tests of writing ability. Dumbing down of tests in this way is required because machines are not yet capable of accurately assessing the types of complex writing we expect of our students. This dumbing down of tests in order to meet the machines present capabilities is akin to the narrowing of the curriculum that was a disastrous result of NCLB. Most importantly, much like the mystery that surrounds the data used for the analysis presented in the widely publicized book *Academically Adrift* that we critiqued in Chapter One, Shermis and Hamner's data is also a closely guarded secret. In situations where research results are used to inform practices as important as the implementation and assessment of the Common Core State Standards, we should insist upon this data being made readily available to other researchers for analysis so that it can be examined closely and debated in the field.

We have said elsewhere and say again that we are not staunch opponents of standardized tests, but we do oppose their current use as primary indicators of student learning as well as faculty achievement. Further, we *are* staunch opponents of dumbing down curricula to meet the limitations of standardized tests and the dumbing down of standardized tests so that they can be scored by machines. If CCSS is to have a chance at succeeding in raising the bar for writing

across the curriculum, then we must insist on practices that lead to embracing the complexity of the CCSS, not altering them in the name of cheap and easy tests. Of course, the original intent of the CCSS is exactly the opposite: To provide national standards that present a robust picture of student achievement by setting ambitious but achievable goals. The CCSS itself, in fact, includes performance-based tasks for writing across the curriculum, tasks that would be difficult to leave to a machine-grader and nearly impossible to shrink to a one-paragraph response.

Standardized tests of writing do not have to lead to a narrowing of the curriculum or to writing assignments designed to meet the limitations of machine-graded scoring. If we look at the sample performance tasks, we see rich writing prompts across the curriculum that can be used to measure depth and breadth of knowledge as well as writing. For example, the CCSS suggest the following as a sample performance task for English Language Arts information texts for grades 6–8[1]:

> Students *trace* the line of *argument* in Winston Churchill's "Blood, Toil, Tears, and Sweat" address to Parliament and *evaluate* his specific claims and opinions *in the text*, *distinguishing* which *claims* are supported by *facts*, *reasons, and evidence*, and which *are not*. (93)

For students in grades 9–10, a sample performance indicator for fiction, poetry, and drama suggests this prompt:

> Students *analyze how* the Japanese film maker Akira Kurosawa in his film *Throne of Blood draws on and transforms Shakespeare's play Macbeth* in order to develop a similar plot set in feudal Japan. (121–22).

The CCSS also support writing across the curriculum, as can be seen in the prompt for grades 9–12 in history/social sciences, sciences, mathematics, and technical subjects:

> Students *cite specific textual evidence* from Annie J. Cannon's "Classifying the Stars" to *support their analysis* of the scientific importance of the discovery that light is composed of many colors. Students *include in their analysis precise details* from the text (such as Cannon's repeated use of the image of the rainbow) to buttress their explanation. (138)

These performance tasks from the CCSS require students to demonstrate complex mastery of literacy—examining argumentative claims, working across

genres and disciplines, and doing so in both reading and writing. To reduce these performance tasks to machine-gradable, short-answer summative assessments would largely undermine the laudable goals of CCSS.

EMPOWERING AND ENGAGED PROFESSIONAL DEVELOPMENT

In working to understand how to improve teacher quality, the Department of Education's Equity and Excellence Commission examined how other nations ensure teaching quality. Their report points out that unlike the United States, teacher training in high-performing countries is based on engagement with common instructional materials that support high-level national standards (22). In *Professional Learning in the Learning Profession: A Status Report on Teacher Development in the U.S. and Abroad*, Linda Darling-Hammond, Ruth Chung Wei, Alethea Andree, Nikole Richardson, and Stelios Orphanos discuss some of these differences:

> In most European and Asian countries, instruction takes up less than half of a teacher's working time (NCTAF, 1996, and OECD, 2007). The rest of teachers' working time—generally about 15 to 20 hours per week—is spent on tasks related to teaching like preparing lessons, marking papers, meeting with students and parents, and working with colleagues. Most planning is done in collegial settings, in the context of subject matter departments, grade level teams, or the large teacher rooms where teachers' desks are located to facilitate collective work.
>
> By contrast, U.S. teachers generally have from 3 to 5 hours a week for lesson planning, usually scheduled independently rather than jointly with colleagues (NCTAF, 1996). U.S. teachers also average far more net teaching time in direct contact with students (1,080 hours per year) than any other member of the Organization for Economic Cooperation and Development (OECD). By comparison, the OECD average is only 803 hours per year for primary schools and 664 hours per year for upper secondary schools (OECD, 2007). U.S. teachers spent about 80% of their total working time teaching students as compared to about 60% for teachers in these other nations, who thus have much more time to plan and learn together, and to develop high-quality curriculum and instruction. (20)

Parts of this description ring true for tenure-track professors in the United States as many of us lack the common spaces and institutional support for ongoing, systematic professional development of our teaching. More troubling is that when it comes to non-tenure-track faculty, many of whom teach five or more writing classes each semester, this is an increasingly apt description of their working conditions. In fact, we would not be surprised to learn that most college writing instructors spend more than 85 percent of their time teaching students and even less time than their K–12 counterparts engaged in professional development. Thus, in our discussion below we abstract beyond K–12 classrooms to include structural changes needed at the college level as well.

As the Equity and Excellence Commission asserts: "Professional development must be embedded in the work day, deepen and broaden teacher knowledge, be rooted in best practice, allow for collaborative efforts, be aligned to the Common Core State Standards and provide the supports, time and resources to enable teachers to master new content, pedagogy and learning tools and incorporate them in their practice" (23). Of course, this general advice must be put into practice in ways that lead to improved student learning.

We usually equate improved student learning with improved teacher quality. And while this equation isn't false, the steps we follow in establishing this equation too often are. For example, as Carrie Leana, a professor of organizational management at the University of Pittsburgh, explains, we've come to believe that the keys to reforming our schools are identifying the most high-achieving teachers and using them as models that others should emulate, hiring outside consultants or identifying coaches, positioning principals as instructional leaders who, "in the language of business, . . . is a line manager expected to be a visible presence in the classroom, ensuring that teachers are doing their jobs." The problem with these beliefs is that while sometimes they can be helpful, there is considerable research showing that this approach alone is relatively ineffective. Leana's research provides evidence that if student learning is to show marked improvement *that is sustainable over time*, "schools must also foster what sociologists label 'social capital'—the patterns of interactions among teachers."

A growing body of research on effective professional development reveals that it is most often rooted in strong teacher networks with high levels of social capital. Indeed, as we argue for writing instruction that is positioned horizontally across the curriculum and vertically through grade levels, we know that this can only be successful in schools where structures exist that support high degrees of social capital among teachers. As Leana argues, when we look at a teacher's social capital, we are asking not only what does an individual teacher know that leads to her success, but also, how does she know it. In other words, how has she gained this knowledge? Where does she turn for new knowledge

and advice when faced with new situations? The research of Leana and her colleagues shows that:

> When a teacher needs information or advice about how to do her job more effectively, she goes to other teachers. She turns far less frequently to the experts and is even less likely to talk to her principal. Further, when the relationships among teachers in a school are characterized by high trust and frequent interaction—that is, when social capital is strong—student achievement scores improve.

For example, in a study of 1,000 fourth and fifth grade teachers from 130 elementary schools in New York City, Leana and her colleagues wanted to find out if social capital is a significant predictor of student gains in math. Their results revealed that students showed higher gains in math when their teachers had higher levels of social capital, that is, "If a teacher's social capital was just one standard deviation higher than the average, her students' math scores increased by 5.7 percent."

Leana's study, as well as others, verifies a practice that permeates much of our professional literature, although often from an anecdotal perspective. That is, what may matter most is the networks that teachers build, formally or informally, in support of professional development and improved student learning. It is these same types of networks writing faculty must build in order to become positive sponsors of literacy, helping to shape the trade routes along which it travels. Formal communities like Bread Loaf, National Writing Project, Teaching and Learning Network, Learning Forward, ReadWriteThink, and the National Council of Teachers of English and their state affiliates can provide teachers with opportunities for professional development that are more than a collection of "Monday morning" worksheets. These venues provide teachers the opportunity to write, to learn, and to participate as active teacher scholars in their own professional development through both local and national networks. Time must be made in the work day of writing teachers to engage in these professional development opportunities.

Likewise, teachers must assert their role in the shared governance of our educational institutions—helping to make decisions with administrators about teaching and learning initiatives. As Gary Olson reminds us: "True shared governance attempts to balance maximum participation in decision making with clear accountability. . . . Genuine shared governance gives voice (but not necessarily ultimate authority) to concerns common to all constituencies as well as to issues unique to specific groups." But shared governance can only be effective if lines

of communication are clear and open, again highlighting the need to make time for such activities during the work day. And, as we expand these opportunities we must engage in research that can determine the most successful structures for sustainable professional development. For example, one of the things we want to know as teacher-scholars is if teachers can gain social capital from both distant and immediate networks, if there are differences in what can be gained, if our professional conferences can do a better job at supporting these networks, and what role collective empirical research can play in increased engagement and improved learning across the curriculum.

While the work of Leana and her colleagues is somewhat unique in its focus on social capital, viewed from another perspective it is also simply one more significant piece of the growing body of research showing the value of effective professional development to improved student learning and achievement. For example, in Linda Darling-Hammond's et al. report on the status of professional development in the United States and high-achieving countries around the world, the researchers conducted a meta-analysis of 1,300 research studies and evaluation reports, and concluded that student achievement can be significantly improved through substantial professional development that ranges from 30–100 total hours, and is spread over six to twelve months: "intensive professional development efforts that offered an average of 49 hours in a year boosted student achievement by approximately 21 percentile points. Other efforts that involved a limited amount of professional development (ranging from 5 to 14 hours in total) showed no statistically significant effect on student learning" (9).

Despite these overwhelmingly positive results, when it comes to professional development, the United States lags far behind our high-achieving counterparts. For example, echoing the results of the research conducted by Leana and her colleagues, the authors of this report also did not find strong support for the effectiveness of coaches or hiring outside consultants in improving student learning and achievement. Further, among some of the most striking findings is that "Nationally, only 17 percent of [K–12] teachers reported a great deal of cooperative effort among staff members, and only 14 percent agreed that they had made conscious efforts to coordinate the content of courses" (25). While congeniality may be found in many schools, true collegiality is rare and can be difficult to sustain (Mindich and Lieberman). And, the difficultly in sustaining this type of work is in large part because in the United States teachers generally have three to five hours per week for tasks related to teaching, such as planning lessons, and this is most often done in isolation. In contrast, teachers in other countries, including high-achieving countries, allow for fifteen to twenty hours per week on tasks related to teaching including working with colleagues (Darling-Hammond, et al.). As Dan Mindich and Ann Lieberman make clear,

collegiality is the cornerstone of professional development. They differentiate congeniality from collegiality by explaining that while congenial relationships are amiable they are often also conflict and risk-averse (18). "Collegial cultures on the other hand develop bonds of trust [and] provide a forum for reflection and honest feedback, for challenging disagreement and for accepting responsibility without assigning blame" (18). Building collegial cultures takes time our teachers are seldom given.

When it comes specifically to literacy, in "What Teachers Need," Darling-Hammond tells us that research presented by the National Center for Literacy Education (NCLE) confirms: "77% of educators, principals, and librarians agreed that developing student literacy is one of the most important responsibilities they have." Despite this finding, the same research reveals that only 32 percent of the respondents frequently create lessons together or reflect on whether or not a lesson worked, only twenty-one percent have time to review student work with each other on a frequent basis, only fourteen percent frequently receive formal feedback from peers, and only ten percent observe the teaching of others on a frequent basis. During the busy workaday of the school week and, it would seem, even at mandatory "School Improvement Days," little time is dedicated to reflection and collaboration among teachers—especially the types of collegial cultures that can lead to improved practices.

How might we develop collegial cultures that further empower and engage professional development? A multi-year study, "Professional Development in the United States: Trends and Challenges," published by Learning Forward and the Stanford Center for Opportunity Policy in Education, and funded by the Bill and Melinda Gates Foundation, is leading the way in answering this question. We have referenced reports from this study in a few sections of this book. At this point we focus on the seven standards for professional learning communities that grew out of the work of Dan Mindich and Ann Leiberman for this study. Professional learning that increases educator effectiveness and results for all students:

1. Occurs within learning communities committed to continuous improvement, collective responsibility, and goal alignment
2. Requires skillful leaders who develop capacity, advocate, and create support systems for professional learning
3. Requires prioritizing, monitoring, and coordinating resources for educator learning
4. Uses a variety of sources and types of student, educator, and system data to plan, assess, and evaluate professional learning

5. Integrates theories, research, and models of human learning to achieve its intended outcomes
6. Applies research on change and sustains support for implementation of professional learning for long-term change
7. Aligns its outcomes with educator performance and student curriculum standards

Exactly how a professional learning community might be operationalized following these standards is dependent on our specific contexts, and as our research enters its next phase, we look forward to identifying and developing local models for writing teachers.

Once in place, just what should empowered and engaged professional development opportunities position us to achieve as sponsors of literacy? We believe that they will allow us to do just what research on best practices, our individual experiences, and professional organizations such as the NCTE and CCCC believe we should be doing. They will lead us to a fuller realization of the many policy statements our professional organizations issue. The current pace of technology and disruptive forces in education mean we are faced with new tools, new genres and subgenres, and new models of composing at a sometimes dizzying speed. We must both document and research these changes while simultaneously teaching them. The timing has never been more crucial for inventing a new responsive and effective writing curriculum in K–graduate school—one that once again places writing, and teachers of writing, in the role of agent. We can't do it alone. It must happen through an essential broadening of our networks and strengthening of our social capital both locally and nationally. This book opens the dialogue for such a movement.

NOTE

1. We have chosen representative examples from the CCSS. Similar examples can be found across all grade levels, K–12, varying appropriately, of course, in complexity and expectation by grade level.