

## Theorizing Failure in US Writing Assessments

---

Asao B. Inoue

California State University, Fresno

*How do teachers define failure when learning to write? We don't ask the question often enough. In this article, I attempt to offer a definition and critique of the nature and production of failure in writing classrooms and programs. I argue that the production of failure in writing assessments can create more purposeful consequences, particularly for those historically most likely to suffer "failures" in writing classrooms: students of color, multilingual students, and working-class students. Drawing upon survey and grade data from California State University, Fresno, I examine two kinds of failure produced in writing classrooms, quality-failure and labor-failure. I argue that quality-failure (associated with judging the quality of drafts) is the least useful kind of failure for writing classrooms, while labor-failure (associated with work and effort) offers better consequences for student-writers and can help articulate a more robust writing construct by including noncognitive dimensions of writing. I conclude by proposing "productive failure" as a future possibility for writing classrooms.*

Failure in writing classrooms may very well be one of the most important yet undertheorized concepts in composition studies. Past discussions about remediation and basic writing and writers have questioned the nature and production of failure by questioning who the basic writer is (Bartholomae, 1987; Hull & Rose, 1989), the inherent racism in basic writing programs and concepts (Fox, 1993, 1999; Jones, 1993), and the relationship between the kinds of languages used by students (often marked by culture, class, gender, and race) and dominant, White, middle-class, academic discourse (Horner & Lu, 1999). In fact, Shaughnessy (1977) has revealed logics to errors in student writing, thus suggesting students' "failure" to write the expected dominant English was a result of "social inequities, not personal failings" (Otte & Mlynarczyk, 2010, p. 8). But in her pedagogy Shaughnessy ignores the politics, the uneven power arrangements occurring around competing languages in the writing classroom. Lu (1991) critiques Shaughnessy's pedagogy, saying it exhibits a "politics of linguistic innocence" (pp. 27, 30) that fails to address the reality that changing students' sentences to match the dominant code of the classroom may change students' ideas and thinking. Taking a macro view, Soliday (2002) convincingly argues that institutional crises from open enrollments in the 1960s created the alleged need for remedial students and classes (pp. 60, 62). Thus institutional needs for remediation constructed the basic writer and, in the process,

writing failure. In effect, Soliday can be used to critique the production of failure in writing programs since she reveals the ways institutions produce remedial students (failing students) for particular historical reasons and institutional needs. Educators produce failure because it suggests something to audiences at their institutions and outside of them, something about the rigor of writing programs, about standards held (against many students), and about teachers doing their job right.

Dressman, Wilder, and Connor (2005) illustrate that understanding the nature of failure, however, is not a simple matter of choosing which theoretical orientation works best—a cognitivist (e.g., Shaughnessy, 1977), sociocultural (e.g., Bartholomae, 1987; Hull & Rose, 1989; Lu, 1991), or macrostructural one (e.g., Horner & Trimbur, 2002; Prendergast, 2003; Trimbur et al., 2011)—nor is it a matter of blending the best of each theory to create a “unified theory” of failure (Dressman et al., 2005, p. 17). Cognitivist perspectives tend to focus on individual learners and their cognitive processes, seeing failure as a product of either pedagogy or individual learners’ behaviors (Dressman et al., 2005, pp. 10–11). This makes literacy practices, such as reading and writing, normative and failure in these practices abnormal. For example, in her discussion of the failure of “decentered pedagogies,” Judy Segal identifies “blaming the student” as an early theory in understanding failure (1996, p. 178). Sociocultural perspectives emphasize the subjectivity of the learner as a “gendered, ethnic, economic user of language” (p. 11), conceiving failure not as abnormal but as a product of mismatched literacy practices. Writing failure stems from irreconcilable differences between expectations of White, middle-class literacies in school and the raced, cultured, classed, and gendered home literacies that learners attempt to use in school. Macrostructural perspectives are similar to sociocultural perspectives, only they focus on “law and history that name specific historical events and legal precedents” as constructors of literacy norms, and thus historical events, laws, and legal decisions create the conditions that produce failure in schools (p. 14).

Failure, in this view, is not just an individual, political, cultural, institutional, or social problem to solve. Failure is a complex blend of these various elements, an inevitable aspect of any writing classroom or program. More important, failure includes noncognitive elements, not identified in the discussions above. Noncognitive dimensions—those associated with things like grit, persistence, openness, or responsibility—are an important part of writing development, as illustrated in the “Framework for Success in Postsecondary Writing” (Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project, 2011, pp. 4–5). Noncognitive dimensions of students’ work are just as influential to their success in school as the cognitive. The dominant articulation of noncognitive dimensions related to success and failure in school by the field of educational psychology is the “five-factor model of personality,” which posits five factors in human personality: “neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness” (O’Connor & Paunonen, 2007, p. 973). Of the five factors, O’Connor and Paunonen have found conscientiousness, interpreted often as “motivation,” to be the most influential to academic success

as measured by meta-analyses of correlations conducted across 23 previously published studies (p. 976). Willingham, Pollack, and Lewis (2002) have found that noncognitive factors also correlate most strongly with student success in a number of measures (i.e., The National Education Longitudinal Study and grades) (p. 19). They conclude that the noncognitive dimension of “scholastic engagement” most affects academic achievement in school (p. 26). It is clear that to understand the nature and production of failure in learning to write, researchers and educators will need to recognize that failure will likely be a product of noncognitive and cognitive dimensions in acts of writing.

In a recent article summarizing his 30-year career, Hayes (2012) suggests that noncognitive dimensions are important to account for in any model of the writing process and thus in any robust writing construct. Hayes’s new model, a revision of the Hayes-Flower model (1980), includes the noncognitive by incorporating “motivation” at the “control level” (Hayes, 2012, p. 371). After citing several studies, he concludes that “whether people write, how long they write, and how much they attend to the quality of what they write will depend on their motivation” (p. 373). This means the noncognitive behaviors involved in intrinsic motivation (writing because the student understands writing’s inherent value) are vital to learning to write effectively. This is akin to “conscientiousness” (motivation) in the five-factor model (O’Connor & Pajunen, 2007), which begs the question: Is extrinsic motivation too artificial to provide enough “control” to aid in learning to write effectively? Elliot and Klobucar (2013) also demonstrate a writing construct that accounts for both the cognitive and noncognitive dimensions of writing. They use the WPA outcomes statement (cognitive; Council of Writing Program Administrators, 1999) and the “Framework for Success in Postsecondary Writing” (noncognitive; Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project, 2011) to produce a “unified model for validation” of their writing assessment decisions (Elliot & Klobucar, 2013, p. 21).<sup>1</sup>

Though these authors have begun to propose models that account for both cognitive and noncognitive dimensions of writing, there has yet to be a discussion of how failure is conceptualized and produced within such models. Doing this work can help overturn pernicious assumptions that failure in writing is abnormal; outside of White, middle-class language expectations; associated with the dispositions of the “remedial” student; or even unwanted in the writing classroom.

In this article, I attempt to do just this work, redefining the nature and production of failure in writing assessment. I discuss ways to understand the nature and production of failure, offer two ways to define failure that affect its production, and (in the process) illustrate the consequences of failure. I ultimately propose a third kind of failure for the writing classroom, one that I believe is possible, even necessary. My purpose is to argue for productive and educative conceptions of failure so that the production of failure in writing assessments also creates more purposeful and pedagogical consequences, particularly for students who historically are closest to failure in writing classrooms—that is, students of color, multilingual students, and working-class students.

## Failure in Writing Assessment

So what does it mean to fail a writing assessment? Failure could be thought of as not succeeding, missing the mark, or receiving a “below average” grade or score. Often “below average” is a grade below a C, or one or two standard deviations below the mean score on a bell curve of all scores. This first definition might be equated with failure produced by norm-referenced tests that compare an individual’s performance with that of the population taking the test. Failure is produced by competition, by how students stack up against each other. It’s a slide-rule model. A student’s writing is measured against a changing, sliding standard, determined by how others do around him or her.<sup>2</sup> Failure might also be thought of as simply not demonstrating achievement of particular learning goals set by the program, teacher, or student, and judged in an assessment. This kind of failure is associated with “standards,” and is often produced by criterion-referenced assessments, ones that provide scores or grades that tell you how well you’ve done on the dimension they are testing but do not compare your score with others’ scores or calculate your score based on comparisons with other scores. It’s a yardstick model. A student’s writing is measured against a fixed yardstick. Failure is produced by comparisons with a fixed ideal of writing, usually Standardized Edited American Academic English (SEAAE). The bottom line is, failure in writing, no matter how you define it, is constructed by writing assessments, and is not simply a result of bad or lazy students, teachers, or pedagogies.

Both of these ways of defining failure in learning to write in a classroom rely on someone (usually the teacher) to judge and determine the markers of failure in writing or on writing tasks. Someone must use the slide rule or yardstick. Thus the nature of failure is co-constructed by at least three elements in any writing assessment: the judge; the system of codes and the calculations used to determine failure in the assessment (e.g., grades, rubrics, ratings, etc.); and the markers in a text or performance that represent the students’ writing efforts. The lynchpin in this kind of failure production is the judge, the teacher; yet many have shown how teachers and other readers are not reliable in scoring or grading (Diederich, 1974; Dulek & Shelby, 1981; Huot & O’Neill, 2009, p. 3; Starch & Elliot, 1912), something Belanoff (1991) explains in her four “myths of assessment.” Williams (1981) similarly argues that when reading student writing, teachers often assume (read for) error. Additionally, Anson (2000) argues that the construction, by the teacher, of a student’s persona in student writing is an important factor in the way the text is judged, and thus the ideological position that a student holds in a text can affect whether the teacher sees an error or deems it important enough to notice (p. 10). Failure becomes messily intertwined with students’ and teachers’ conflicting ideological stances as much as with teachers’ phenomenological reading practices that may presuppose error.

Another complication in defining what it means to fail in writing is the phenomenon of the natural diversity of teacher judgments on writing brought about by the nature of assessment as a system in schools, institutions, and society, and by how those structures produce products (like grades and SAT scores) that become

naturalized. This systemic reality makes all writing assessment both sociocultural and macrostructural. Grades typically represent the degree of a student's failure or success, but this also means that when grades are a product of an assessment, they produce degrees of failure outside of the student. Arguments against grading as a practice in writing classrooms are not new (e.g., Bleich, 1997; Bernard-Donals, 1998; Elbow, 1993, 1997). Many have offered alternatives to conventional grading in writing classrooms (Allison et al., 1997; Elbow, 1997; Inoue, 2005; Zak & Weaver, 1998). Few propose to eliminate the final course grade; however, the literature on grading contracts in writing studies does argue to eliminate grades on assignments (e.g., Danielewicz & Elbow, 2009; Reichert, 2003; Shor, 2009; Smith, 1999). The point I'm making is that many agree that without grades on writing, failure looks and is experienced by students very differently: Can one fail when writing a personal diary entry? Even if teachers can be certain about what constitutes a failing performance (even if they can be reliable judges of writing), is it always necessary to designate certain performances as failures? As a sociocultural and macrostructural system, must writing assessment always construct some kind, or the same kind, of failure?

These questions are more important when we consider that writing assessments as systems create failure in writing or students by producing the very characteristics they purport to measure (Hanson, 1993, p. 284). Just as Soliday (2002) and others have shown, teachers and institutions have remedial students because tests identify remedial students through their failure to write in dominant ways. This occurrence hinges on the way tests tend to turn around "the relation between the signifier and signified" (Hansen, p. 287). This inversion means assessment results, like SAT scores, become more significant than what they are supposed to signify—that is, learning or writing ability. For teachers who desire to help students learn to write, a focus on the signifier of grades becomes a learning-to-write problem. This focus gets in the way of student learning, as many writing theorists and others have already argued (e.g., Bleich, 1997; Diederich, 1974; Elbow, 1993, 1997). Perhaps the most well-known discussion of this phenomenon is Gould's (1981) case concerning the reification of IQ test scores and the testing of intelligence, which has had clear racist consequences (p. 24). When people act as if the signifier of an IQ score is something real, a reality that is signified in the world and that has been measured, they forget that the IQ score is produced by the test and does not actually exist before the test. The same naturalizing and reifying happens with the use of grades and failure through writing assessments. Writing assessments, then, produce failure as a product in the system (such as the "remedial" designation), reifying it and then naturalizing failure in student populations.

Horner and Trimbur (2002) critique a "tacit policy of unidirectional monolingualism" that influences writing courses and pedagogies (p. 595), identifying a "chain of reifications" used to "locate individual learners on a sequence of development" in writing classrooms and programs. While they don't say it explicitly, Horner and Trimbur locate the problem that these reifications create as a writing assessment problem, and thus a problem of the possible overproduction of failure

through writing assessments. In short, it is not that teachers and writing programs have students who do not write SEAAE; it is that they have not clearly identified the boundaries of failure in writing. Instead, they've left it as anything other than a reified version of a local SEAAE.

Volosinov (1929), Vico (2002), and contemporary linguists (e.g., Greenfield, 2011; Lippi-Green, 1997) make similar arguments about the nature of language. Volosinov (1986), for example, argues that language cannot truly be understood synchronically, because it changes constantly and there is “no inert system of self-identical norms.” Instead, language consists of a “ceaseless generation of language norms,” and thus is “a ceaseless flow of becoming” (p. 66), in large part because an “utterance is a social phenomenon” (p. 82). When teachers and programs reify language practices in writing assessments by using only a local SEAAE to impose a standard, they work against language's inherently dynamic nature. Failure, in these cases, might just be identifying when a language practice has moved on, upgraded, or evolved.

Teachers who must assess writing and administrators of programs that use such writing assessments likely realize this problem to some degree and live with the fiction that they can capture the slick eel of SEAAE with little harm done to students. But even if they are okay living with the fiction, there still exists an assessment problem that unevenly affects social, ethnic, and racial populations other than the local dominant White one. Their local SEAAE, as Horner and Trimbur explain, is used to “locate individual learners on a sequence of development fixed in its order, direction, and sociopolitical significance” (2002, p. 595). For multilingual students and many students of color, this means an automatic placement at the bottom of the order. Historically and linguistically, multilingual students and many students of color are closest to failure, and as the literature on remediation and basic writing has illustrated (Horner & Lu, 1999; Lu, 1992; Soliday, 2002), the body of color is usually associated with remediation, and thus with failure. Similarly, Matsuda (2006) takes the argument a step further to demonstrate a heightened version of this problem with multilingual and international student populations, arguing that US writing instruction tends to assume a native English speaker as the default writing student and that this assumption is “out of sync with contemporary sociolinguistic reality” (p. 641). I agree, and add that this assumption further reifies multilingual students and their writing as failed writing from the start. When looking for bad writing or failure, teachers don't just find it because tests construct it or, perhaps, overconstruct it. They find it in particular places, students, and kinds of writing because those places, people, and texts are already reified as failure.

But failure isn't just constructed in the judgments of teachers or in the codes and expectations (e.g., SEAAE) used to determine grades and rankings in assessments. In fact, failing grades wouldn't mean much if they weren't used in some way beyond the classroom. Failure is constructed through the ways educational institutions, employers, and others outside of schools use the products of their assessments—the codes, numbers, grades, and other marks that signify student

performance and ability. Grades as products of assessments become signifiers for other things that have no connection to the reasons for any particular assessment of writing. Through others' uses, grades signify so many things about students and their capabilities that they quickly become *the most important thing* in school, *the thing* students end up caring most about. In one sense, the reification of failure through grades transforms intrinsically motivated learning into extrinsic motivators, grades, and scores.

Clearly, the uses of assessment products that identify failure and success have psychological effects on students, which have a bearing on how they approach learning. These psychological effects illuminate the concept of failure as something that structures human activities and behavior. For instance, they illuminate failure's relationship to noncognitive dimensions of students. Carr (2013) discusses the negative affective dimensions of failure in the writing classroom, considering how failure in writing is often felt as if it were a moral failure. In contemporary psychological studies, the effects of failure in an assessment system are well understood. Pulfrey, Buchs, and Butera (2011) demonstrate that grades encourage students to engage in "performance-avoidance," or "avoiding normative incompetence," rather than focusing on "performance-approach" goals, or "attaining normative competence" in an activity (p. 683). Performance-approach goals are essentially behaviors that are "instigated or directed by a positive or desirable event or possibility," whereas performance-avoidance is when "behavior is instigated or directed by a negative or undesirable event or possibility" (Elliot, 1999, p. 170). In short, performance-avoidance is behavior that attempts to avoid failure, but not necessarily pursue success or learning.

These two kinds of behaviors have a long history in the psychological research on motivation and test anxiety, which I will touch on briefly. McClelland (1951), for example, has theorized two similar kinds of achievement motivations that individuals use: a negative kind that focuses on avoiding failure and a positive kind that focuses on the possibility of achieving success. Covington and Beery (1976) further describe classrooms as having either an orientation toward success or toward failure, suggesting that students tend to adapt to these orientations by cultivating either a "success orientation" or a "failure-avoidant" one (Elliot, 1999, p. 171). Pulfrey et al. (2011) have found that "the anticipation of a grade for an activity increased pre-task performance-avoidance goals when compared with a non-graded activity, an effect that did not manifest itself in the case of performance-approach goals" (p. 688).

The bottom line is that writing assessments that produce failure, or that have possible failure present in the system (but not necessarily experienced by a student), have psychological consequences for all students in the system that negatively affect their learning. As previous research on both the association of noncognitive dimensions with academic writing success (Elliot & Klobucar, 2013; Willingham et al., 2002) and the importance of "conscientiousness" (motivation) in the five-factor model to academic achievement (O'Connor & Paunonen, 2007) suggests, failure in the system structures students' behaviors toward learning. It may even

overdetermine those behaviors, because the future uses of a signifier of failure, like an *F* grade, have severe, negative consequences. The need to avoid failure is psychologically and cognitively very strong.

If (a) judgment and grades on writing are unreliable and always biased, (b) the nature of assessment systems is such that failure is created, structured, reified, and naturalized as something to expect from the assessment, which signifies many other things beyond learning, and (c) external uses of the products of our assessments foster anxiety that moves students toward performance-avoidance behaviors that are counter to learning, then it becomes clear that failure is a complex systemic phenomenon with structural, social, affective, cognitive, and noncognitive dimensions. Failure is not simply a product of bad or lazy students or bad or lazy teachers. Failure is not simply located in student texts, students themselves, teachers' judgments, or the processes, codes, artifacts, or products that circulate in writing assessments (e.g., rubrics, feedback, and grades). Failure is defined and produced through the interaction of all these elements in a writing assessment system. For without the assessment, there is no failure, even if there are performances. But all failure isn't the same. That is, I'm not arguing to eliminate failure in writing assessments, or that we can. I'm arguing that given what we know about its nature and production, we might consciously construct it differently.

In what follows, I illustrate that possibility by using my own writing program as one site in which failure has been consciously reconstructed. I use student survey and grade data to demonstrate the pedagogical value of distinguishing between quality-failure and labor-failure. I then suggest the implications of this distinction for the field of writing studies to consider what I call "productive failure."

### **Fresno State's First-Year Writing Program**

To better contextualize the data I use in order to illustrate the two conceptions of failure below, it is necessary to describe the university, writing program, exit survey, and grade data from which I draw. Fresno State is a part of the California State University system, the largest university system in the US. A four-year, public university, Fresno State is a designated Historically Hispanic Serving Institution that enrolls roughly 22,500 students each year. In fall 2012, Fresno State students were 38.8% Hispanic (Mexican), 28.8% White, 14.6% Asian (mostly Hmong, and Chinese), and 4.4% African American.<sup>3</sup> Matching these demographics, the writing program enrolls approximately 1,350 students annually.

In 2007, Fresno State's writing program installed a directed self-placement (DSP) model based on Grand Valley State University's (e.g., Royer & Gilles, 1998). Within this model, students choose three course options to complete their writing requirement.<sup>4</sup> The program's curriculum is based on entering academic conversations, and its longest standing text is Graff and Birkenstein's (2009). In fall 2009, Academic Literacy 1, the course that half of all students choose as their first of two in the writing program, was changed to a credit/no-credit course with a mandated use of a grading contract. All courses in the writing program also require a port-

folio (with universal guidelines), which is used in some significant way in course assessments and grading.<sup>5</sup>

Additionally, the program conducts annual, anonymous, online exit surveys of all its students, in which approximately 30%–45% of students voluntarily participate each year.<sup>6</sup> Surveys are administered by email in the last week of instruction of the student's final writing course and are sent by an automated online service. Students have two to three weeks to complete the survey before it is closed. The survey service produces graphs showing numbers and percentages of student responses by program (not course section) and allows for disaggregation of the data by race and gender, which students self-identify in survey responses. The survey data I look at below are from fall 2009 to fall 2012. All grades that I discuss below were gathered by the Office of Institutional Effectiveness, which keeps track of such data at Fresno State. I use these grade data as one preliminary measure of failure in learning to write in writing courses, even though I realize course grades are by no means a definitive measure of learning to write.

## Nature and Production of Failure at Fresno State

### *Quality-Failure*

There are at least two ways writing assessments define the nature of failure. The first, which I call quality-failure, exists in all conventionally graded courses that use judgments of quality to determine success and failure in writing. I am tempted to use the term *grade-failure*, since writing assessment systems using quality-failure almost always denote that failure in terms of grades (the *D* or *F* grade signals failure). But quality-failure can also be nongraded. Take, for instance, a scenario in which a teacher is responding to (not grading) a draft, but nonetheless requires a student to revise or turn in a new draft because the first one is not sufficient or does not meet some criteria. This failure can be quality-related but not graded. I'm also tempted to use the label *SEAAE-failure*, since as I've discussed already, "quality" is usually associated or synonymous with a local SEAAE, but this term isn't as immediately intuitive as *quality-failure*. Thus, when I say "quality," I really mean judgments against or comparisons to some local SEAAE.

As Table 1 demonstrates, course failure in Fresno State's writing program was greater in the conventionally graded courses that were offered during AY 2005–06 (the final year all writing courses used conventional writing assessment systems). In subsequent years, after a contract grading system was introduced, failure occurred less frequently across racial populations.<sup>7</sup> If, as I have argued above, quality-failure is inherent in conventional grading systems, then more course failure was produced by quality-failure. In 2005–06, the production of quality-failure was at its highest, and course failure was produced more often in African American, Hmong, and Latino/Latina student populations (see the first row in Table 1).<sup>8</sup> At Fresno State, quality-failure was associated more strongly with students of color than with White students.

The bottom line is that courses using grades and judgments of quality on writing produced more frequent failure in particular racial populations at Fresno State

**TABLE 1. Data on Course Failure in Fresno State’s First-Year Writing Program (English 1 and English 5B)**

Academic Year	African American			Asian American (Hmong)			Latino/Latina			White		
	N	No. FAILED	% FAILED	N	No. FAILED	% FAILED	N	No. FAILED	% FAILED	N	No. FAILED	% FAILED
2005–06 (ENGL 1)	198	45	22.7%	454	90	19.8%	843	188	22.3%	788	121	15.4%
2009–10 (ENGL 5B)	130	25	19.2%	158	16	10.1%	682	75	11.0%	292	21	7.2%
2010–11 (ENGL 5B)	109	18	16.5%	195	19	9.7%	685	65	9.5%	273	23	8.4%
2011–12 (ENGL 5B)	66	11	16.7%	160	16	10.0%	553	78	14.1%	158	10	6.3%

than courses using grading contracts. The old writing courses (Engl 1) produced primarily quality-failure, with White students always producing the smallest proportion of quality-failure and African American students the largest (however, not substantially more than Latino/Latina students in English 1). The consequences of quality-failure can be seen better when juxtaposed with the second kind of failure.

**Labor-Failure**

The second kind of failure, labor-failure, is often associated with not achieving or demonstrating a defined degree of effort, quantity of written products, and/or amount of time spent on an activity such as reading or drafting. In the contract-graded courses at Fresno State, these criteria are the primary ways course grades are produced. Labor-failure is associated with noncognitive dimensions such as conscientiousness, persistence, and motivation. Labor-failure is most often judged on a binary scale: work is done or not done, okay or not okay. Labor-failure is found in conventionally graded and alternatively graded courses; however, if quality-failure exists in a writing assessment, typically labor-failure will not, since they contradict one another. Labor-failure, because it defines failure to write as not spending enough time or producing enough writing on an assignment, is more likely to be seen by students as fair. Everyone can spend an hour on a draft or write a page. When students don’t, they understand this failure is not a personal judgment about them or their ideas; rather, their failure is an identification of what they have done (or have not done, but could have). However, failure to meet labor requirements in a course that uses labor as a way to grade could be seen as unfair if students perceive the requirements as too much work (Spidell & Thelin, 2006).

To illustrate how the nature and production of labor-failure are different from those of quality-failure, consider Fresno State’s contract-graded writing courses. In the exit surveys from 2009–12, 83% (912 of 1,099 surveyed) of Academic Literacy II (or Engl 5B in the above table) students reported being contract-graded, while 56.7% (416 of 734 surveyed) of students enrolled in the one-semester course option of Accelerated Academic Literacy (also known as Engl 10) reported the same. Thus the majority of instances of course failure produced in Academic Literacy II

(and at least half in Accelerated Academic Literacy) were produced by labor-failure in courses. While it's hard to know for sure what happens in classrooms or what causes changes in failure rates, when compared with the previous English 1 course, the numbers in Academic Literacy II suggest that a focus on labor-failure generally reduces the amount of course failure, with a notable exception. Table 1 shows the amount of labor-failure produced in Academic Literacy II (Engl 5B courses) in the same three-year period (2009–12), comparing it with the labor-failure produced in English 1 (2005–06). While the production of labor-failure lowers all course failure rates, African American students still produced an unusually high amount of failure in Academic Literacy II, but not as much as in English 1, which used quality-failure to designate course failure. However, considering the small number of African American students in the program, perhaps their slightly higher failure rates were affected by social and/or linguistic isolation. And if Fresno State writing teachers are anything like their counterparts at the University of California, Merced (just north on Highway 99), then this finding is curious, because Fowler and Ochsner (2012) found that UC Merced writing teachers did not penalize student writing that showed Latino/Latina or African American Vernacular English textual markers (p. 123).

Comparing Academic Literacy II with English 1 shows, on average, 8.5% fewer instances of course failure by racial population in all three academic years listed in Table 1. Perhaps most interesting about these rates is the consistency of labor-failure rates in Academic Literacy II, and how Hmong students have benefited most (relatively speaking), reducing their course failure from 19.8% in English 1 to an average of 9.9% in Academic Literacy II for academic years 2009–12. A similar drop in failure rates occurred in the Latino/Latina population, moving from 22.3% in English 1 to an average of 11.5% in Academic Literacy II. Meanwhile, Whites averaged about an 8.1% reduction in course failure, going from 15.4% in English 1 to an average of 7.3% in Academic Literacy II. Why these changes happened is hard to say for sure, but the apparent fact that labor-failure encourages and rewards noncognitive dimensions in writing might be one cause, and it may suggest that a good portion of failure in writing courses has more to do with noncognitive aspects of student work than the cognitive dimensions of writing on which teachers usually focus. A closer look at what is happening in courses and portfolios may offer some insight. I'm not suggesting that the presence and production of labor-failure causes the drops in course failure (I've done no statistical analyses to make such claims), but these two factors do appear to be associated.

To be clear, in the data (2009–12), labor-failure is produced by a student not meeting grading contract obligations (e.g., page and word counts, turning in assignments on time, etc.). Among the African American racial population, an average of 17.5% failed courses for the academic years 2009–12. As indicated in Tables 2 and 3, the program's exit surveys show that, on average, 67.8% of the African American racial population found grading contracts to be "effective" or "very effective," and 62% said they were "happy" or "very happy" with the grading contract. Meanwhile, in blind portfolio readings done by the program in academic

years 2008–10, the average overall ratings given to African American students’ portfolios were 3.73 (spring 2009) and 3.27 (spring 2010) on a 6-point scale (where 3 was the lower boundary for passing quality) (Inoue, 2012a, p. 88).<sup>9</sup> That is, these students produced proficient portfolios, on average. Labor-failure in writing courses at Fresno State creates the conditions for students to meet cognitive expectations of a local SEAAE writing construct, even though failure is determined mainly by noncognitive dimensions, and the same students tend to be mostly okay with how failure is produced. Thus, the incorporation of labor-failure as the primary way to fail courses and assignments makes for a robust, if not fairer, construct of writing for program and classroom writing assessments.

TABLE 2. Grading Contract Effectiveness Rates for English 5B<sup>a</sup>

Academic Year	African American			Asian American (Hmong)			Latino/Latina			White		
	N	EFFEC-TIVE NO. <sup>b</sup>	EFFEC-TIVE % <sup>c</sup>	N	EFFEC-TIVE NO.	EFFEC-TIVE %	N	EFFEC-TIVE NO.	EFFEC-TIVE %	N	EFFEC-TIVE NO.	EFFEC-TIVE %
2009–10 (ENGL 5B)	40	37	92.5%	37	30	81.1%	174	123	70.7%	90	69	76.7%
2010–11 (ENGL 5B)	13	9	69.2%	80	60	75.0%	136	101	74.3%	45	31	68.9%
2011–12 (ENGL 5B)	12	5	41.7%	69	53	76.8%	106	84	79.2%	30	23	76.7%

<sup>a</sup>As indicated by student responses to exit surveys.

<sup>b</sup>Number of students who reported that the grading contract was “effective.”

<sup>c</sup>Percentage of students who reported that the grading contract was “effective.”

<sup>d</sup>Exit surveys were not conducted in AY 2005–06.

TABLE 3. Grading Contract Happiness Rates for English 5B<sup>a</sup>

Academic Year	African American			Asian American (Hmong)			Latino/Latina			White		
	N	HAPPY NO. <sup>b</sup>	HAPPY % <sup>c</sup>	N	HAPPY NO.	HAPPY %	N	HAPPY NO.	HAPPY %	N	HAPPY NO.	HAPPY %
2009–10 (ENGL 5B)	40	30	75.0%	37	26	70.3%	171	124	72.5%	93	66	71.0%
2010–11 (ENGL 5B)	13	9	69.2%	79	61	77.2%	137	100	73.0%	45	32	71.1%
2011–12 (ENGL 5B)	12	5	41.7%	68	45	66.2%	105	75	71.4%	31	21	67.7%

<sup>a</sup>As indicated by student responses to exit surveys.

<sup>b</sup>Number of students who reported that they were “happy” with the grading contract.

<sup>c</sup>Percentage of students who reported that they were “happy” with the grading contract.

<sup>d</sup>Exit surveys were not conducted in AY 2005–06.

It should be noted that the writing program changed in a few important ways in the summer of 2011, which may have affected these data. The program's vision, goals, and outcomes did not change, but a new codirector replaced me as codirector and a revised curriculum was introduced, along with a revised grading contract that was slimmer but articulated labor-failure in the same ways as the older contract. There was less explicit training on the grading contract for new and returning teachers (mostly teaching assistants). It could be argued that the old, longer version of the contract (2,816 words) taught teachers more how to use the contract than the newer, slimmer one (975 words). These changes could have caused some disruption in course failure, and might suggest why there were some differences in survey results in 2011–12, particularly for African American students. In 2011–12, grading contract effectiveness and happiness rates of African American students dropped the most of all racial populations (see Tables 2 and 3). As I have shown elsewhere (Inoue, 2009a, 2012a, 2012b), students' reactions to contract grading often fluctuate by local student racial populations, so it is reasonable to expect that various racial populations will respond differently to the presence and production of labor-failure in courses.

Sociocultural and macrostructural reasons may explain African Americans' tempered response to labor-failure, and these reasons may have created non-cognitive behaviors that showed up in surveys. For instance, the dramatic drop in reported effectiveness of contracts seen only in African American students (see Table 2) may be a result of a desire to learn the local SEAAE for social and economic mobility, which may affect African American students' expectations to learn SEAAE when they enter college. When quality-failure is absent in a course but expected, the labor-failure may be perceived as less effective in helping with learning to write. A similar explanation can be made of the lower happiness rates of African American students in the same year (see Table 3). In the year when the contract was arguably least explained, and when teachers were least trained on it, African American students (a vulnerable racial population at Fresno State because of their smaller numbers) may have reacted more strongly against the presence of labor-failure, yet it did not change their general production of grade-failure, nor did it cause more quality-failure in final portfolios. This bodes well for labor-failure.

### ***Criticisms of Labor-Failure***

One criticism of producing labor-failure could be its lack of attention to quality. What about the outcomes of a labor-failure assessment system? How do I know student writing will still get better in my class? How do I know my standards are still high enough? How do I make sure my students will produce good enough writing if I'm not judging quality? How do I know I'm preparing them for the future? I have two responses to these potential criticisms, and they take very different paths. The first is that labor-failure does not necessarily mean lowering standards. At Fresno State, grade distributions are similar to what they were before contracts—we give a few more A's now program-wide, but about the same relative amounts of B's and C's (see Table 4)—and this pattern stays consistent across all racial populations, with African American students the main exception. Labor-failure produced in the

**TABLE 4. Grade distributions of all students for Fresno State's English 1 and English 5B**

Academic Year	<i>n</i>	A	B	C	Fail
2005–06 (ENGL 1)	2,568	31.3%	33.2%	16.3%	16.2%
2008–09 (ENGL 5B)	1,671	46.5%	31.5%	11.8%	10.2%
2009–10 (ENGL 5B)	1,493	36.0%	39.9%	13.1%	11.1%
2010–11 (ENGL 5B)	1,491	48.6%	30.4%	12.1%	8.9%

Note: Data were provided by the Office of Institutional Effectiveness at Fresno State.

program did not reduce the quality of writing in program portfolios assessed by blind readings, which were rated at “proficient” or better levels in all racial populations, with all students—including African American students (Inoue, 2012a, pp. 89–90)—showing growth in most measured outcomes of the program (p. 88). In fact, the Asian Pacific Islander population (mostly Hmong) showed statistically significant higher average ratings on one key program outcome, “summary and conversation” (pp. 86–87), which is a demonstration of these students’ use of sources and ability to enter academic conversations.

You may be wondering how it is possible to maintain quality, as measured cognitively based on a local SEAAE, even though most teachers are not grading quality in classrooms. Simple. Students find reasons to learn and grow as writers when their labor is truly honored, and they listen more carefully to feedback when grades are out of the way, perhaps especially because their writing labor is being acknowledged and quality is assumed to be a consequence of that hard labor. As researchers studying the positive association of noncognitive dimensions with college success explain, “cognitive ability reflects what an individual *can* do, personality traits [noncognitive dimensions] reflect what an individual *will* do” (O’Connor & Paunonen, 2007, p. 972).

My second response is that focusing on outcomes, like how close any student is to a local SEAAE, disregards the most powerful rhetorical aspects of writing as an epistemological process, as a process of labor that produces often unexpected, meaningful consequences. Writing is a generative and surprising labor, at least when it’s done well. In his critique of outcomes assessment, Gallagher (2012) advocates moving from thinking in terms of outcomes to thinking in terms of “consequences” (p. 47). He suggests that thinking of students’ writing in terms of outcomes narrows our view (p. 48), and theirs. It reduces possibilities. Teachers miss a lot when they only look for evidence of outcomes in drafts. They miss the expected and unexpected consequences *in* and *from* the labors of writers. They miss what can be learned from sweaty work but not seen in a product or outcome. They miss helping students’ with noncognitive behaviors, and rewarding them for those behaviors.

Seeing what students do in courses as mostly a question of outcomes, as cognitive dimensions in written texts, tends to frame the products of student labor as

failure if they do not match a local SEAAE. What gets obscured is the consequences (good, bad, or ambiguous) of much of the labor that multilingual students and students of color struggle through when reading and writing. While Gallagher's argument that "aims" are a better way to orient ourselves toward the "intended and unintended results of our interactions with students" (p. 47) is quite persuasive, it's not enough. We have to start trusting our students. I find that most teachers' resistance to writing assessment systems that produce labor-failure are rooted in a distrust in students, a lack of faith in students and their abilities to find intrinsic motivation to do this most human endeavor (to communicate and create with language), or too much faith in what teachers' practices produce, too much faith in what judging, marking, and grading on the quality of students' writing do for their students.

### **Redefining Failure in Writing Assessment**

When considering these two kinds of failure and my responses to criticisms of labor-failure as the preferred of the two for writing assessments, it helps to examine more carefully how judgments are made about writing, and this will lead me to suggest productive failure. Haswell (1998) explains three ways of making judgments in the context of writing placement systems, but these methods of judgment operate in any writing assessment system where readers read student writing and make judgments from those readings, producing quality-failure. In brief, the three categorization methods are (a) "classical," categorizing writing by how close or similar a particular student text is to an idealized construct of writing (a set of qualities, features, or attributes) (p. 245); (b) "prototype," categorizing by comparing the present text with an idealized text (a prototype) in the reader's mind (p. 246); and (c) "exemplar," categorizing by comparing an array of firmly held, real examples of ideal writing (exemplars) in the reader's memory with the present text (p. 247).

If a teacher is using classical categorization, say in combination with a rubric, to evaluate and grade essays, then quality-failure is produced by her perception of a piece of writing that doesn't contain the right amount of idealized, static features that constitute the writing assignment (i.e., a set of textual abstractions). If she uses prototype categorization, quality-failure is produced by her judging an essay as not meeting some threshold of similarity to an idealized construction of the writing assignment in her mind (i.e., an ideal text). If the teacher uses exemplar categorization, then quality-failure is produced by her perception of writing that is farthest from some threshold of similarity to an assemblage of intact memories of real, similar kinds of writing (i.e., past, real texts).

Who fails and how often can be quite different in each method. How many prototypes or exemplars of good writing for any given assignment does a White, middle-class teacher have for, say, a local Black English Vernacular or Spanglish, or an English native of Hong Kong? How would a teacher design writing assessments so that such Englishes could achieve something other than failure? One could argue that these are not the Englishes that students aim to learn in writing courses,

and therefore these ways of judging aren't flawed in this manner. Failure simply reveals when writing doesn't meet a minimum standard of competency. But I am not talking about success, outcomes, or the aims of a writing class. I am talking about how students fail to achieve those outcomes and aims, or more accurately how our writing assessment systems produce quality-failure through particular ways of reading in those settings.

Let me be clear: I'm arguing that there are two kinds of outcomes to define in the writing we judge, the possible ways to succeed and the possible ways to fail. The absence of perceived "quality" shouldn't necessarily mean failure. Leaving quality-failure as simply the absence of quality (a local SEAAE), or as not fitting closely enough to the categories, prototypes, or exemplars we hold as ideal writing, will likely doom many students of color and multilingual students to failure automatically. In producing quality-failure in our assessment systems, writing teachers and administrators need to consider the ratio that exists between the conceivable instances of success and failure in our judgment processes, and those instances should take into account the particular students who inhabit classrooms, so that both students and teachers can control quality-failure and not let it simply happen.

Sociocultural and macrostructural problems further trouble Haswell's three judgment methods. When situating classrooms and writing assessments in larger societal environments that have few published examples of writers of color or multilingual writers writing in their home Englishes or meshed versions of them, judging by prototype and exemplar can be dubious, since teachers will rely on the prototypes and exemplars most at their disposal, which are usually White, middle-class ones (i.e., SEAAEs). In prototype and exemplar readings, it seems to matter a great deal who judges writing and what experiences with language those teachers have. The classical method isn't much better. It tends to assume a false, static, idealized notion of language that many scholars I've cited tell us simply does not match the way language dynamically exists in the world. If, as Volosinov (1986) claims, language is a "ceaseless flow of becoming," then writing that defies our expectations might be tomorrow's conventions and normal discourse, the cutting-edge of today's literacy, or at least a place for critical insight into the local SEAAE and Englishes that clash in the contact zones of classrooms.

It's arguable, however, that assessing writing only in terms of labor, effort, or quantity reorients students toward the work and behaviors most teachers hope students will learn, and thus toward performance-approach behaviors, noncognitive dimensions of academic success. It also avoids the damaging psychological effects, such as performance-avoidance and low self-efficacy, that grading by quality can cause many students, most notably students of color, working-class students, and multilingual students. Most importantly, when a piece of writing does not meet a teacher's sense of a local SEAAE, it does not require a failing grade, since writing quality does not denote labor-failure. Thus, labor-failure is produced only when students do not do enough work. This framework opens up room for translingual and other critical pedagogies that aim to question SEAAEs or broaden the academy.

## **Conclusion: Toward Producing Better Failure in Writing Assessments**

The distinction between quality-failure and labor-failure shown in the Fresno State data points to failure as social and public. It further suggests the need for both K–12 and postsecondary writing teachers, administrators, and scholars to understand failure’s nature and modes of production from an ethical standpoint, as productive failure.

While both quality- and labor-failure denote a judgment that a student has not met some kind of expectation, be it of quality or labor, productive failure is not so easily measured in or designed into a writing assessment. It is a way to re-define quality- or labor-failure by transforming it through the process of a course or curricular sequence, through reseeing what unsuccess means and can become in drafts and work. Productive failure is positive and signals the opportunity to learn, grow, and develop in purposeful ways. Thus productive failure can happen when students and teacher negotiate learning at the point where an absence of success, quality, or sufficient labor seems visible.

Productive failure pushes students and teachers to negotiate and redefine quality- or labor-failure in writing assessments so that failure opens new ways of seeing and languaging. Productive failure takes the question of “quality” in a draft and pushes the teacher and student to investigate, research, and negotiate the expectations of the local SEAAE and the student’s discourse, and even to question what “error” means in each discourse. This process resembles Horner’s (1992) call for students to negotiate error and theorize about it with teachers (p. 188), and agrees with Lu’s (1994) “border pedagogy,” allowing teacher and students to “map the contact zone” in student drafts (p. 447). It also encourages code-meshing by not penalizing it (Martinez & Young, 2011; Young, 2007, 2009), instead making it an opportunity to negotiate, inquire, and discuss language practices as such (Canagarajah, 2013). Perhaps most important, productive failure falls outside the conventional grading system. It does not produce grades or evaluations by teachers; instead, it produces judgments, investigations, negotiations, and discussion among students and with the teacher about expectations, new drafts, and future practices.

Finally, productive failure in writing assessment systems pushes schools, colleges, and universities to expand, to become more inclusive of more kinds of students and their linguistic worlds. Productive failure can signal where, in students’ drafts, linguistic and cultural appropriation, colonization, and difference occur, and thus where more informed and purposeful decisions about those drafts (and the students) can be generated by teachers and students together. It allows daily work to be judged by noncognitive behaviors, ones promoted in the “Framework for Success in Postsecondary Writing” (Council of Writing Program Administrators et al., 2011):

- Putting in time on drafts without considering quality (engagement)
- Willingness to explore new territory, ideas, formats, or ways of understanding (openness and creativity)

- Sustaining interest despite a lack of apparent results (persistence)
- Taking responsibility for one's own learning and that of one's peers (responsibility)

Thus, productive failure may most easily come out of labor-failure, refocusing attention on the noncognitive by encouraging students and teachers to search for and engage with their differences in language, attitudes, culture, and material circumstances, creating productive borderlands and linguistic contact zones.

Two caveats qualify my discussion of writing failure. First, my references to failure in writing assessments and writing courses are an attempt to discuss only *failure when learning to write*, not failure to write or failure of a particular text to do something rhetorical (as judged by a reader, like a teacher), although these issues are related in writing classrooms. The result of failure when learning to write usually equates to course or assignment failure (a failing grade); thus, I've assumed that grades are one indicator of failure (although not definitively or comprehensively so) when writing. Second, some may expect a deeper discussion of the literature on motivation from the field of psychology. While acknowledging this perspective as an important aspect of failure deserving of continuing attention from writing scholars and teachers, I have not exhaustively delved into that literature here. I have framed failure in writing as an issue of assessment rather than motivation, because the latter stance continues to place the problem—and thus most of its solution—in the laps of students. While students' motivations and desires play an important role in their success in writing classrooms, my goal is to pay closer attention to the system creating and producing failure, not to the individuals to whom failure is attached in that system. Thus, my concerns focus on how failure is defined and produced in and through writing assessments.

In many schools, it may be difficult to create a writing assessment system that avoids quality-failure. There are always tensions and contradictions, but there are also always small ways to begin. In light of productive failure, I invite writing teachers and administrators to put aside a host of assumptions and see what happens. Put aside the assumption that teachers are the guardians of "good writing" and that assuming this role will somehow help students write in ways that will improve their future life chances. Put aside the assumption that teachers cannot trust students to do what is asked of them, cannot trust that when asked to write for three hours, students will write for three hours (regardless of what they bring back to the classroom). I invite teachers to put aside the idea that grading or evaluating student writing by some quality standard is more helpful than harmful to students; that if they do not understand a student's text or how he or she arranges it, then his or her writing is flawed or misses the mark; and that a writing classroom that does not penalize nondominant Englishes is somehow not doing its job in a diverse and global world where SEAAEs are hegemonic. Instead, I propose that teachers frame what they experience as a productive clash at a border that they do not, and cannot, control.

Shor (2009) regards classroom grading as a social, public practice that is embedded in the “unrelenting struggle” and conflict of students’ lives (p. 21). Thus it’s safe to say that in Shor’s paradigm, failure is not just a product of the writing assessment system, not just a product of the judgments of teachers or their expectations. Failure is also a product of the social forces in the school and larger society that create that classroom and its material conditions, including the future uses of our grades. Failure is also a product of the social conflicts inside and outside the classroom and in society at large. Oftentimes, failure identifies the subaltern in hegemony, yet failure helps “for[m] us into the people we become” and is inevitable (Shor, p. 21). The question is, what kind of failure do we create, and what are the consequences to our students, ourselves, and the institutions we work in when we produce that kind of failure?

### ACKNOWLEDGMENTS

I thank Xuanning Fu, the interim director of the Office of Institutional Effectiveness, and research analysts Hongtao Yue, Dimitri Rogulkin, and Joann Venturi for their great help in compiling all the data.

### NOTES

1. The “Framework for Success in Postsecondary Writing” was produced by the Council of Writing Program Administrators, National Council of Teachers of English, and National Writing Project (2011) and identifies eight “habits of mind” that contribute to successful writing in college: curiosity, openness, engagement, creativity, persistence, responsibility, and metacognition (pp. 4–5).
2. The traditional slide rule is a device that calculates multiplication and division tables by sliding a middle section of the ruler to align marks on it. Thus, results change as you slide the ruler.
3. All Fresno State demographics come from Fresno State’s Office of Institutional Effectiveness (<http://www.fresnostate.edu/academics/oie/index.html>).
4. The courses that correspond to the three DSP options are: (1) Accelerated Academic English (one semester), (2) Academic Literacy I and II (two semesters), and (3) Advanced English Strategies for Multilingual Speakers and Academic Literacy I and II (three semesters). The majority of students choose between the first two options. Each semester, only a few dozen choose the third.
5. Fresno State’s DSP and program portfolio is discussed in another place (Inoue, 2009a).
6. Enrollment in the writing program at Fresno State has been very consistent. In AY 2011–12, the enrollment by racial population was as follows (enrollment numbers for English 5A, the precursor to English 5B, are not included in the numbers below):

	<b>Engl 5B</b> <b>(Academic Literacy II)</b>	<b>Engl 10</b> <b>(Accelerated Academic Literacy)</b>
African American	66	56
Asian American (Hmong)	240	261
Latino/Latina	553	623
White	158	395

7. Note that I use “failure” as both a count and noncount noun, meaning at some times it is appropriate to identify the amount of failure produced, while at others, it is more appropriate to speak of the degree of failure produced. This is because writing assessments produce both the nature of failure and its quantity.

8. All data on courses, grades, and students were provided by Fresno State’s Office of Institutional Effectiveness.

9. Overall portfolio ratings are averages of five independent ratings (on a scale of 1–6, with 3 or 4 as passing or “proficient” quality) of the following dimensions as demonstrated in each portfolio, which correspond to program outcomes: reflection, reading and writing practices, summary and conversation, community participation (collaboration), language coherence. The methods and sampling for portfolio ratings are discussed in other places (Inoue, 2009a, 2012a).

## REFERENCES

- ANSON, C. (2000). Response and the social construction of error. *Assessing Writing*, 7, 5–21.
- BARTHOLOMAE, D. (1987). Writing on the margins: The concept of literacy in higher education. In T. Enos (Ed.), *A sourcebook for basic writing teachers* (pp. 66–83). New York: Random House.
- BELANOFF, P. (1991). The myths of assessment. *Journal of Basic Writing*, 10(1), 54–66.
- BERNARD-DONALS, M. (1998). Peter Elbow and the cynical subject. In F. Zak & C. C. Weaver (Eds.), *The theory and practice of grading writing: Problems and possibilities* (pp. 53–65). Albany: State University of New York Press.
- BLEICH, D. (1997). What can be done about grading? In L. Allison, L. Bryant, & M. Hourigan (Eds.), *Grading in the post-process classroom: From theory to practice* (pp. 15–35). Portsmouth, NH: Boynton/Cook.
- CANAGARAJAH, A. S. (2013). Negotiating translingual literacy: An enactment. *Research in the Teaching of English*, 48(1), 40–67.
- CARR, A. (2013). In support of failure. *Composition Forum*, 27(Spring). Retrieved from <http://www.compositionforum.com>
- COUNCIL OF WRITING PROGRAM ADMINISTRATORS. (1999). WPA outcomes statement for first-year composition. *WPA: Writing Program Administration*, 23(1–2), 59–66.
- COUNCIL OF WRITING PROGRAM ADMINISTRATORS, NATIONAL COUNCIL OF TEACHERS OF ENGLISH, & NATIONAL WRITING PROJECT. (2011). *Framework for success in postsecondary writing*. Retrieved from <http://wpaacouncil.org/framework>
- COVINGTON, M., & BEERY, R. (1976). *Self-worth and school learning*. New York: Holt, Rinehart, & Winston.
- DANIELEWICZ, J., & ELBOW, P. (2009). A unilateral grading contract to improve learning and teaching. *College Composition and Communication*, 61(2), 244–268.
- DIEDERICH, P. (1974). *Measuring growth in English*. Urbana, IL: National Council of Teachers of English.
- DRESSMAN, M., WILDER, P., & CONNOR, J. J. (2005). Theories of failure and the failure of theories: A cognitive/sociocultural/macros-structural study of eight struggling students. *Research in the Teaching of English*, 40, 8–61.
- DULEK, R., & SHELBY, A. (1981). Varying evaluative criteria: A factor in differential grading. *The Journal of Business Communication*, 18(2), 41–50.
- ELBOW, P. (1993). Ranking, evaluating, and liking: Sorting out three forms of judgment. *College English*, 55(2), 187–206.
- ELBOW, P. (1997). Taking time out from grading and evaluating while working in a conventional system. *Assessing Writing*, 4(1), 5–27.

- ELLIOT, A. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169–189.
- ELLIOT, N., & KLOBUCAR, A. (2013). Automated essay evaluation and the teaching of writing. In M. D. Sheremis & J. Burstein (Eds.), *Handbook for automated essay evaluation: Current applications and new directions* (pp. 16–35). New York: Routledge.
- FOWLER, J., & OCHSNER, R. (2012). Evaluating essays across institutional boundaries: Teacher attitudes toward dialect, race, and writing. In A. B. Inoue & M. Poe (Eds.), *Race and writing assessment* (pp. 111–126). New York: Lang.
- FOX, T. (1993). Standards and access. *Journal of Basic Writing*, 12(1), 37–45.
- FOX, T. (1999). *Defending Access: A Critique of Standards in Higher Education*. Portsmouth, NH: Boynton/Cook.
- GALLAGHER, C. (2012). The trouble with outcomes: Pragmatic inquiry and educational aims. *College English*, 75(1), 42–60.
- GOULD, S. J. (1981). *The mismeasure of man*. New York: Norton.
- GREENFIELD, L. (2011). The “Standard English” fairy tale: A rhetorical analysis of racist pedagogies and commonplace assumptions about language diversity. In L. Greenfield and K. Rowan (Eds.), *Writing centers and the new racism* (pp. 33–60). Logan: Utah State University Press.
- GUBA, E. G., & LINCOLN, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: SAGE.
- HANSON, F. (1993). *Testing testing: Social consequences of the examined life*. Berkeley: University of California Press.
- HASWELL, R. H. (1998). Rubrics, prototypes, and exemplars: Categorization theory and systems of writing placement. *Assessing Writing*, 5(2), 231–268.
- HAYES, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
- HORNER, B. (1992). Rethinking the “sociality” of error: Teaching editing as negotiation. *Rhetoric Review*, 11(1), 172–199.
- HORNER, B. (1996). Discoursing basic writing. *College Composition and Communication*, 47(2), 199–222.
- HORNER, B., & LU, M. (1999). *Representing the “other”: Basic writers and the teaching of basic writing*. Urbana, IL: National Council of Teachers of English.
- HORNER, B., LU, M., ROYSTER, J. J., & TRIMBUR, J. (2011). Language difference in writing: Toward a translingual approach. *College English*, 73(3), 303–321.
- HORNER, B., & TRIMBUR, J. (2002). English only and U.S. college composition. *College Composition and Communication*, 53(4), 594–630.
- HULL, G., & ROSE, M. (1989). Rethinking remediation: Toward a social-cognitive understanding of problematic reading and writing. *Written Communication*, 6(2), 139–154.
- HUOT, B. (2002). *(Re)articulating writing assessment for teaching and learning*. Logan: Utah State University Press.
- HUOT, B., & O’NEILL, P. (2009). *Assessing writing: A critical sourcebook*. New York: Bedford/St. Martin’s.
- INOUE, A. (2005). Community-based assessment pedagogy. *Assessing Writing*, 9(3), 208–238.
- INOUE, A. (2009a). Self-assessment as programmatic center: The first year writing program and its assessment at California State University, Fresno. *Composition Forum*, 20(Summer). Retrieved from <http://www.compositionforum.com>
- INOUE, A. (2009b). The technology of writing assessment and racial validity. In C. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 97–120). Hershey, NY: IGI Global.
- INOUE, A. B. (2012a). Grading contracts: Assessing their effectiveness on different racial

formations. In A. B. Inoue & M. Poe (Eds.), *Race and writing assessment* (pp. 79–94). New York: Lang.

INOUE, A. (2012b). Racial methodologies for composition studies: Reflecting on theories of race in writing assessment research. In L. Nickoson & M. P. Sheridan (Eds.), *Writing studies research in practice: Methods and methodologies* (pp. 125–139). Carbondale: Southern Illinois University Press.

JONES, W. (1993). Basic writing: Pushing against racism. *Journal of Basic Writing*, 12(1), 72–80.

LIPPI-GREEN, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.

LU, M. (1991). Redefining the legacy of Mina Shaughnessy: A critique of the politics of linguistic innocence. *Journal of Basic Writing*, 10(1), 26–40.

LU, M. (1992). Conflict and struggle: The enemies or preconditions of basic writing? *College English*, 54(8), 887–913.

LU, M. (1994). Professing multiculturalism: The politics of style in the contact zone. *College Composition and Communication*, 45(4), 442–458.

MARTINEZ, A., & YOUNG, V. A. (Eds.). (2011). *Code-meshing as World English: Policy, pedagogy, performance*. Urbana, IL: National Council of Teachers of English.

MATSUDA, P. K. (2006). The myth of linguistic homogeneity in U.S. college composition. *College English*, 68(6), 637–651.

MCCLELLAND, D. (1951). *Personality*. Chicago: Dryden.

O'CONNOR, M. C., & PAUNONEN, S. V. (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43, 971–990.

OTTE, G., & MIYNARCZYK, R. W. (2010). *Basic writing*. West Lafayette, IN: Parlor Press.

PRENDERGAST, C. (2003). *Literacy and racial justice: The politics of learning after Brown v.*

*Board of Education*. Carbondale: Southern Illinois University Press.

PULFREY, C., BUCHS, C., & BUTERA, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology*, 103(3), 683–700.

REICHERT, N. (2003). Practice makes perfect: Contracting quantity and quality. *Teaching English in the Two-Year College*, 31(1), 60–68.

ROSCH, E. (1978). Principles of categorization. In E. Rosch & B. R. Lloyd (Eds.), *Cognition and categorization* (pp. 26–48). Hillsdale, NJ: Erlbaum.

ROYER, D. J., & GILLES, R. (1998). Directed self-placement: An attitude of orientation. *College Composition and Communication*, 50(1), 54–70.

SEGAL, J. (1996). Pedagogies of decentering and a discourse of failure. *Rhetoric Review*, 15(1), 174–191.

SHAUGHNESSY, M. (1977). *Errors and expectations: A guide for the teacher of basic writing*. New York: Oxford University Press.

SHOR, I. (2009). Critical pedagogy is too big to fail. *Journal of Basic Writing (CUNY)*, 28(2), 6–27.

SMITH, J. A. (1999). Contracting English composition: It only sounds like an illness. *Teaching English in the Two-Year College*, 26(4), 427–430.

SOLIDAY, M. (2002). *The politics of remediation: Institutional and student needs in higher education*. Pittsburgh: University of Pittsburgh Press.

SPIDELL, C., & THELIN, W. H. (2006). Not ready to let go: A study of resistance to grading contracts. *Composition Studies*, 34(1), 35–68.

STARCH, D., & ELLIOTT, E. (1912). Reliability of the grading of high school work in English. *School Review*, 20, 442–457.

VICO, G. (2002). *Vico: The first new science* (L. Pompa, Trans.). Cambridge, United King-

dom: Cambridge University Press. (Original work published 1725)

VOLOSHINOV, V. (1986). *Marxism and the philosophy of language* (L. Matejka & I. R. Titunik, Trans.). Cambridge, MA: Harvard University Press. (Original work published 1929)

WILLIAMS, J. (1981). The phenomenology of error. *College Composition and Communication*, 32(2), 152–168.

WILLINGHAM, W. W., POLLACK, J. M., & LEWIS, C. (2002). Grades and test scores: Account-

ing for observed differences. *Journal of Educational Measurement*, 39(1), 1–37.

YOUNG, V. A. (2007). *Your average nigga: Performing race, literacy and masculinity*. Detroit: Wayne State University Press.

YOUNG, V. A. (2011). “Nah, we straight”: An argument against code switching. *JAC*, 29(1–2), 49–76.

ZAK, F., & WEAVER, C. C. (EDS.). (1998). *Theory and practice of grading writing: Problems and possibilities*. Albany: State University of New York Press.

**Asao B. Inoue** is an associate professor at California State University, Fresno, and the special assistant to the provost for Writing Across the Curriculum.

*Initial submission:* April 2, 2013

*Final revision submitted:* September 30, 2013

*Accepted:* October 2, 2013

### Call for Nominations: Richard Meade Award

The Conference on English Education is now accepting nominations for the Richard Meade Award for Research in English Education. This award recognizes published research-based work that promotes English/language arts teacher development at any educational level, of any scope, and in any setting. Such work may be published in any of the many modalities that reflect the intellectual and theoretical work in the field.

Criteria for the award are as follows: (1) the selection committee may consider published material of any length and modality focused on the preparation and education of preservice and/or practicing English language arts teachers; (2) eligibility extends to any research-based approach that promotes English language arts teacher development at any educational level; and (3) to be considered, studies must have been published (traditionally or digitally) between January 1, 2013, and December 31, 2013.

Nominations may be made by any English language arts educator or by self-nomination. Nominations for studies published between January 1, 2013, and December 31, 2013, must be received no later than **May 1, 2014**.

Send nominations to CEE Meade Award, NCTE, 1111 W. Kenyon Road, Urbana, IL 61801-1010 or [cee@ncte.org](mailto:cee@ncte.org), Attn: CEE Administrative Liaison. Winners will be notified in July 2014 and announced at the 2014 NCTE Annual Convention in Washington, DC.