

CHAPTER 7.

HOW EFFECTIVE CAN LABOR-BASED GRADING CONTRACTS BE?

Any conscientious teacher will want to know just how effective can labor-based grading contracts be in a writing or English course? This last chapter attempts to answer this question, and “effectiveness,” as I hope you will see, is a deceptively complex concept to determine. There are a number of kinds of data one might measure in order to make arguments of how effective any writing assessment ecology is. One might use final course grades, judgments of final course portfolios, student perceptions of learning, the amount of output or writing done in the course, student attitudes about writing or their own writing at the end of the course, how well students do in later writing-intensive courses (determined by course grades, authentic writing from those courses, or other external measures), retention rates in the university after the course, how students use or employ writing as a practice after the class, among numerous other things. The point is: There are many ways to measure effectiveness, and the data one collects and analyzes to come to some conclusions about the effectiveness of a classroom’s assessment ecology should match the assessment ecology’s goals and primary purposes. Put another way, how any teacher (or program) finds out how effective grading contracts are will be dictated by what they believe evidence for effectiveness of contracts are in the classroom, or what the teacher believes effectiveness means. So in some ways, this question is deceptively difficult to answer. Asking the question seriously, should allow a teacher or WPA to consider what exactly their definition of effectiveness of an ecology can or should be, and what measures or data they might use to make arguments about effectiveness.

As will be clear in my discussion in this chapter, one cannot separate a labor-based grading contract from the rest of the classroom’s assessment ecology, so I tend to use the two terms synonymously, speaking of labor-based grading contracts or contract ecologies as synonymous. This means that this chapter is really about understanding the effectiveness of labor-based grading contract ecologies as a whole. As the previous chapters attest to, one cannot simply insert a labor-based grading contract in a course that is otherwise no different from a conventionally graded course. In all assessment ecologies, there are many moving parts. They are complex systems (Inoue, *Antiracist* 86-92).

It should be clear that the goals of a course’s assessment ecology can be different from its course learning goals. Course learning goals may be used to un-

derstand how effective a course is at achieving a number of curricular things that will vary from program to program. The goals I speak of in this chapter are best used to understand how well a labor-based assessment ecology is working, if one accepts them as I've discussed them in this book. Additionally, as you'll see, some or all of these ecological goals could be used as course learning goals too, but I am speaking of them in this chapter as assessment ecology goals, or the goals I have for determining the effectiveness of a labor-based assessment ecology.

Therefore, in this chapter, I begin by briefly discussing what effectiveness is for an assessment ecology, then why I use the noncognitive literature to help me formulate the five ecological goals that I then discuss as ways to understand a labor-based grading contract ecology's effectiveness. When discussing each goal, I attempt to offer a rationale for the ecological goal as one that can help define effectiveness in some way for a labor-based grading contract ecology, then offer several ways to measure that goal (possible evidence from the ecology), which come from my own ecologies. I mean for this chapter to be useful to both individual writing teachers and writing program administrators tasked with assessing effectiveness of writing programs. I do not think it is necessary that all labor-based grading contract ecologies have all five goals I discuss, but they are ones I use to help me collect evidence and make arguments (often to myself) about effectiveness.

UNDERSTANDING EFFECTIVENESS FOR LABOR-BASED GRADING CONTRACT ECOLOGIES

What does it mean to say that a labor-based grading contract ecology is effective? Do grade distributions explain effectiveness in such ecologies? Do ratings on portfolios explain effectiveness? Do students' opinions about writing explain effectiveness? Do students' confidence levels about themselves as writers explain effectiveness? Does the amount of labor in reading and writing tasks during the semester or quarter explain effectiveness? Any of these kinds of evidence can be compelling, however, I caution teachers and administrators. If your primary ways of seeing the effectiveness of labor-based grading contract ecologies is by measuring how close students' writing is to a predetermined, dominant Discourse, a standard, that is, to a White racial *habitus*, you haven't understood why labor-based grading contracts are important for all students, nor why I find them important for socially just, equitable, and inclusive writing classrooms.

In "Grading Contracts: Assessing Their Effectiveness on Different Racial Formations," I offer three measures of effectiveness of labor-based grading contracts that I use to understand the different degrees of effectiveness contracts had on different racial formations in my writing program at California State Univer-

sity, Fresno, which also situated those courses in a directed self-placement program. They are: “(1) the quantity of work produced, (2) the quality of writing produced in class, and (3) student reactions to and acceptance of the contract itself” (82). I formulated this three-pronged approach from the literature on grading contracts in writing studies (discussed in Chapter 2), which is limited and has no actual studies of effectiveness. I still find these three measures useful for external audiences as an argument for the effectiveness of many labor-based grading contract ecologies, either at the program or classroom level. Using such a three-pronged approach to understanding a contract ecology’s effectiveness may not fit every writing course. More important, since that study, I’ve found other more nuanced ways that account for effectiveness just as well or better.

The above three kinds of evidence do not tell you how to answer the question: *What makes my labor-based grading contract ecology effective for the students I work with in my classroom?* What does effectiveness mean in your kind of assessment ecology with your students and your course’s specific goals? Knowing the terrain of possible answers to this question can help a teacher decide whether a labor-based grading contract system is right for them, and the kinds of evidence necessary to explain its effectiveness to themselves, their students, and others. Effectiveness for labor-based grading contract ecologies, like all assessment ecologies, is determined by the specific learning goals the course attempts to promote. That is, an assessment ecology is only effective if it can be argued it achieves the goals for which it was designed.

A writing teacher should have good reasons for grading the way they do, or using grades the way they do, and those reasons should be evidence-based, and continually checked. Why do you give the grades you do? Why do you use the methods you use to produce course grades? Why do you think those methods help the students in front of you? How do those methods for grading help students achieve the course’s learning goals? What evidence do you have to help you answer these questions? The point is, your goals for using labor-based grading contracts, like any grading ecology, should be explicit so that you know what kind of evidence will help you understand how effective the assessment ecology is at doing what it was designed to do. If half your students fail, it may be that your grading ecology is doing exactly what it is designed to do, even though failing half your class may not be one of your own goals. Because labor-based grading contracts circulate labor in the ecology as a way to calculate course grades, assessing the effectiveness of them makes a teacher have to confront the multiple ways that dominant White *habitus* already regulate all of our typical outcomes, standards, and assumptions about what makes a grading practice effective.

Too often, the presence of a dominant White *habitus* is the measure of effectiveness of an assessment ecology. That is, when we see a student producing such

writing, we tend to think all is okay with that student. The course is working for them, and by default, so is the assessment ecology. But is it? Is that student working hard, challenging themselves, engaging with the material? Are they doing less than they could, or as much as they should? Are they just going through the motions? Are your methods of grading helping the student accomplish the goals of the course or are your grades simply rewarding them for who they are or what they already can do? Or maybe, your grades are just measuring and ranking students, which will privilege students with a White racial *habitus*, and harm student who do not embody it. Can that kind of ecology be called effective? Can an ecology be called effective if the teacher does not need to be concerned about what students are doing, how they are laboring to learn, and only concerned about the products submitted for a grade?

This leads to other effectiveness questions that may not seem like ones. Is your ecology punishing other students for who they are? Is it punishing students who are other than the ones who embody the ideal *habitus* that your standards and grading practices use to grade so-called quality? Can an assessment ecology be called effective if it punishes, accidentally or not, an expected group of students—that is, if punishment in the ecology is unevenly distributed among the social groups that make up that course?

I do not mean to oversimplify classroom labor dynamics, as Chapter 3 demonstrates. I realize that a teacher should not assume that all or most of their students who come to the classroom embodying a White, middle-class, racial *habitus* are not engaging enough or laboring enough—my discussion of Goal #5 (below) will demonstrate one way that I can see this. I also think it is too simple to say that all students who do not match such *habitus* are hard-working and doing all they can in the course. But I do think that grading by a standard encourages those students who already can reproduce that standard to do just enough to achieve the grade they want, and not much more. There are few rewards or encouragements structured into conventional quality-based grading ecologies that invite students to do more labor, effort, or learning than is necessary for the “A.” An ecology is not effective when this happens pervasively if learning is the goal.

Grades on assignments also do not encourage students to think, feel, or behave in ways that orient them toward learning or growth, a sign of an effective assessment ecology. Grades often encourage a fixed mindset, one that is opposed to a more beneficial growth mindset, as described in psychological studies.⁴⁸ And this

48 Much of the psychological research on growth mindset (Dweck; Dweck et al.) reveals how important and influential a person’s notions about ability are to their development in any domain. For instance, Anindito Aditomo found that university students with higher measurable growth mindset about their academic ability in a challenging Statistics course performed better than their peers when experiencing setbacks and challenges (215-16). In another study, growth

harms those students who can already produce standardized writing by keeping them from really working hard, really learning all they can, just as much as it harms those students who are not rewarded for their extra efforts because they cannot yet produce such standardized writing. Thus effectiveness, as you'll see in the rest of this chapter, has mostly to do with understanding students' laboring and noncognitive domains practiced by students. These are measures of effectiveness because they last longer and are more flexible for their unknowable futures.

As discussed in Chapter 6 (see the question, "To what extent can labor-based grading contracts address learning or course objectives or outcomes?"), I agree with Chris Gallagher's concerns about narrow learning outcomes, preferring more open concepts, such as goals or consequences. I think of the difference between goals and outcomes like this. An outcome is a destination that you want every student to get to in a course, a specific place, which is predetermined, *a priori*, regardless of who ends up in the course or how hard anyone works. This location must be decided beforehand, then judged by someone. And who do you think those people are statistically speaking in the US? What training did they get and what language, what *habitus*, do they need to succeed and be in the position to make such decisions about outcomes in writing? And therein lies the White language supremacist dangers of using outcomes to determine an ecology's effectiveness. One can measure the effectiveness of an outcome by who is judged to be at the specific destination that the outcome describes. In writing courses, this is typically done by judging writing products for their quality, or comparing them to a single White standard encapsulated in the outcome—but really this comparison is one judge's comparison to *their own* conception of the outcome, no matter how it is articulated.

Take the way decisions are made in the National Football League when referees must determine the placement of a ball or whether a foot was out of bounds. The need for instant replay protocols demonstrates the problems with knowing exactly where someone is located in the field of play. It is always a matter of the perspective of the viewer/camera/referee/judge. Ask a differently located teacher about where a particular student is located relative to their notions of a predefined outcome and their vantage point relative to the student, and you get a different answer for each teacher. The back judge and the line judge in football see the playing field differently because they are located in different places on the gridiron, and each sees different players differently because those players are located in different places on the field. This is why the game needs a back judge and a line judge positioned in different places on the field, and not just one judge. But we usually only have one judge in classrooms. In football, where the play ends, and where a foot is located mindset was shown to have a positive effect on academic achievement in low-income students in Chile (Claro et al.).

is not an objective fact, but a subjective judgment based on where the judge is judging, and the game is designed around this fact of subjective judgment by using multiple judges and instant replay. Students' and teachers' linguistic, cultural, racial, gendered, and social locations in the classroom are also differently located, and create different judgments of the same text, and yet we have but one judge, located by necessity in one location, then we ask that judge to determine how effective their classroom was.

And there are many other reasons why such judgments by teachers are not just idiosyncratic and inconsistent but racist and White supremacist in effect. In Chapter 2, I explain Richard Haswell's articulation of categorization theory that psychologists have used to theorize the way holistic judgments are made. Categorization theory reveals why such judgments of student writing are not only a comparison to some fictionalized construction in the judge's head, but that that fictionalized construction is not only idiosyncratic and vulnerable to inconsistency, but will be determined by the White racial *habitus* acquired through White texts and authors that the judge has in their head because of the training that they could get in school. How many Latinx or Black or Native American authors did you read in school, in college, who talked about your disciplinary subject, say writing pedagogy? Furthermore, implicit racial bias (a version of confirmation bias) (Banaji and Greenwald), and stereotype threat (Steele; Steele et al.), discussed in Chapter 4, and Daniel Kahneman's studies of the WYSIATI and availability heuristics our brains use to think fast (and make lots of errors), discussed in Chapter 6 (see, "Don't some students want or need grades?"), each provide further reasons to be very suspicious of trying to measure the effectiveness of learning in a writing classroom by anyone's judgments of other people's writing, particularly when we know that students and teachers embody by necessity different *habitus*.

There are other reasons to be cautious about using a single person's judgments of language to decide grades in diverse classrooms. In *Outliers*, Malcolm Gladwell reviews the research on patterns of aggression historically in communities in Kentucky, as well as the aggression exhibited by many people from the southern US. These studies show that southerners can be more easily turned aggressive than people from other areas of the country. Social scientists Richard Nisbett and Dov Cohen dubbed the source of such aggression, "the culture of honor," which Gladwell cites. Gladwell concludes his chapter on these studies by considering the source of various patterns of judgment and behavior in society. He says, "Cultural legacies are powerful forces. They have deep roots and long lives. They persist generation after generation, virtually intact, even as the economic and social and demographic conditions that spawned them have vanished, and they play a role in directing attitudes and behavior that we cannot make sense of our world without them" (144-45).

He calls these conditions that affect decision making and judgment, “the steady accumulation of advantages,” explaining that “when and where you were born, what your parents did for a living, and what the circumstances of your upbringing all make a significant difference in how well you do in the world” (175). While he avoids such discussions, Gladwell provides a way to see how racial implicit bias and particular ways that WYSIATI and the availability heuristics are a part of larger, racist and White supremacist social patterns in the US, how they will affect the success of students of color in schools, that it matters that just a few generations ago most Blacks were slaves, that just a few generations ago, Whites owned everything, decided everything, including what counted as ideal speech and writing, and that such social patterns are historically durable, in the same ways that Omi and Winant describe racial formations as historically dynamic, in the same ways that Bourdieu discusses the durability of *habitus*.

Not so ironically, in a footnote, Gladwell offers more proof of the socio-cultural, durable inheritance of behavior by illustrating it with language spoken by Appalachians, taken from David Hackett Fischer’s research, concluding that “whatever mechanism passes on speech patterns probably passes on behavioral and emotional patterns as well” (175). If we know anything about language and rhetoric, from Gorgias to Protagoras, from Burke to Perelman and Albrechts-Tyteca, from Lakoff to Vico, it is that our words are epistemological in nature. To understand and make sense of our worlds is to make words about them and share them. With words come attitudes and behaviors, biases and logics. So racism and White supremacy, because they are in our past, in our educational structures, in our words, from Kindergarten to college, and in our academic disciplines’ values and discourses, they—racism and White supremacy—remain with us, even as we denounce them. It is no wonder it is so difficult to escape White language supremacy in schools and society, so hard for Black students or indigenous students to even get to college. Therefore, measuring the effectiveness of writing assessment ecologies by predefined outcomes is the best recipe for White language supremacy in today’s schools, even when teachers, administrators, and others explicitly disavow such a thing.

Learning goals, on the other hand, are broader and describe not a location but a direction, and often the speed or method of traveling in that direction. So exactly where one stops at the end of the semester is unknown, but the direction they are headed and the way in which they travel along that direction can be measured and determined. It is evidence of the direction and methods of traveling that I discuss in this chapter as evidence of effectiveness of labor-based grading contract ecologies. There may still be questions of judgment and perspective, but there will be fewer of them, and they usually are less important in determining how effective the ecology has been for any given student, and have a significantly smaller impact

on any student's opportunities for learning and success in the course. The exact destination of any student in the ecology is not as important as the directions they are headed, and the speeds at and methods by which they travel on the landscape.

DEFINING THE NONCOGNITIVE IN LABOR-BASED GOALS

Because of these reasons, I articulate my course's purpose and learning goals, but not particular learning outcomes. It is the student-determined directions and methods of travel that I'm most interested in understanding when trying to make arguments about the effectiveness of my assessment ecologies, not the specific locations at which my students may end the course. My goals are dictated, then, by the central purpose of my labor-based grading contract ecology. I will not defend this purpose, in part because I think it is self-explanatory, and because I do not hold it up for all writing teachers to adopt, rather I show it so that you can understand how the goals I offer below work toward a larger course purpose that can accommodate individual students' purposes and goals for taking the course. The larger purpose I offer may be articulated as:

The purpose of this writing course is to encourage students to engage in a willingness to labor in mindful and meaningful reading and writing practices that lead them toward an awareness of language (and perhaps its politics) in a compassionate and safe environment that makes the course's opportunities for learning and all grades attainable by all students, no matter where they come from or the version of English they use.

My labor-based grading contract ecologies work from pedagogical assumptions that place noncognitive competencies above cognitive ones, which tend to need an exclusionary standard by which to measure them, among other problems that I've discussed elsewhere (Inoue, *Antiracist*). Equally important, noncognitive competencies feed cognitive ones. To labor in the spirit that the grading contract asks of students, I believe, can build noncognitive competencies, which I explain below. I wish to note that my use of noncognitives as competencies assumes that everyone has the ability to practice and develop what the psychological literature tends to refer to as noncognitive "traits" or "skills." I assume any noncognitive competency or trait is along a continuum and is itself a process that is practiced by all students to some degree. In my discussion below, I refer to noncognitive traits or skills when referring to the literature that uses those terms, but I use noncognitive competency when referring to their application to the labor-based grading contract classroom because "competency," for me, suggests a process-oriented set of practices that everyone has access to.

The literature on noncognitive traits (Bowles and Gintis; Farkas; Gutman and Schoon; Zhou) argues strongly that noncognitive traits are more important to student future success in and outside of school than any cognitive skills students might acquire in college.⁴⁹ Some have even found stronger links between noncognitive traits and school decisions and attendance, as well as wages in the labor market after college, than cognitive ability (Heckman et al. 477-78). Some noncognitive traits (i.e., “conscientiousness” and “future-orientation”) have been shown to determine whether college students attend classes/lectures and spend additional time studying (Delaney et al. 189). Research has also shown how noncognitive traits are associated with cognitive skills and academic outcomes more generally (Bowles and Gintis; Farkas; Heckman et al.; Lleras). Thus, noncognitive competencies appear to be one good way to articulate goals for a writing course, particularly since they map closely to labor by being practices and processes that students do.

In a UNESCO report on noncognitive skills, Kai Zhou explains noncognitive traits as: “the ‘patterns of thought, feelings and behaviours’ (Borghans et al., 2008) that are socially determined and can be developed throughout the lifetime to produce value. Noncognitive skills comprise personal traits, attitudes and motivations” (2). While acknowledging the contested nature of the concept of noncognitive skills, Gutman and Schoon offer this definition of them in learners:

The term “non-cognitive skills” is used to contrast a variety of behaviours, personality characteristics, and attitudes with academic skills, aptitudes, and attainment. The concept was introduced by sociologists Bowles and Gintis (1976) to focus on factors other than those measured by cognitive test scores. They highlighted the role of attitudes, motivation and personality traits, rather than academic skills, as determinants of labour market success. Their findings have been reinforced by more recent studies, which have demonstrated the significant role of non-cognitive skills (i.e., attitudes, motivation and personal characteristics) over and above cognitive skills in shaping labour market outcomes, social behaviour and health (Farkas, 2003, Heckman et al., 2006). (7)

So the practices and processes that noncognitive competencies refer to also have associated with them certain attitudes, feelings, and behaviors. As Gutman and

⁴⁹ Borghans et al. offer a detailed footnote (#3 on page 973) that lists various studies since Bowles and Gintis’ work in the mid-1970s that associate noncognitive traits to earnings. They also summarize much of the research on noncognitive traits’ predictive power in the article.

Schoon's definition suggests, we might understand the difference between cognitive and noncognitive competencies in the same way that we differentiate between deduction and induction, by the direction of their logical movement and the nature of the conclusions made.

In writing courses, cognitive traits are often treated as outcomes or products that students produce. This means the value of student products, and therefore the value of student writing, is determined by cognitive competencies that are predefined, like a major premise in a syllogism, or an assertion (what Aristotle calls an *apophansis*) that a current instance is judged against in order for the syllogism to be valid. In *Prior Analytics* (I.2, 24b18–20),⁵⁰ Aristotle discusses deduction, but it is the nature of the logical movement, a move from a predefined assertion, like an outcome or standard in a writing course that is critical to see here. In a syllogism, then, the conclusion is based on a comparison of a new or present instance of writing to the *a priori* assertion (*apophansis*), a standard, or outcome. The key to the difference is the predetermined, *a priori* nature of the assertion, standard, or outcome. In writing courses, this predetermined assertion is based on cognitive competencies that align with White middle-class *habitus*.

In inductive logic, there is no *a priori* assertion before the present instance being considered. The present instance of, say, writing from a student, is not compared to some predefined outcome or standard to determine its quality or legitimacy. Instead, inductive logic is open-ended, and so are noncognitive competencies. We don't know how something like engagement or persistence will manifest itself, but we do know that these noncognitive competencies manifest in a wide range of practices and outcomes, many of which lead to success in a number of ways. These outcomes are always unique to some degree. What is not unique is their source, the noncognitive competency, the practices, attitudes, and behaviors. Thus, one observes noncognitive competencies more than judges them (against some standard), and this forms the induction. This means that noncognitive competencies, if used as goals in a writing course, will not tell us what students will produce exactly, as much as indicate the competencies students can practice activating in the course to produce what they do. Thus students' abilities to meet particular goals (or directions) are the best way to determine effectiveness of labor-based grading contract ecologies, and noncognitive competencies have shown to be a good, well-researched way to articulate such goals.

The most influential articulation of noncognitive traits is the "Big Five," which are: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (or emotional stability) (Costa and McCrae 258; Borghans et al.;

50 A good overview of Aristotle's logic and how he defines and discusses deduction is offered by Robin Smith in the *Stanford Encyclopedia of Philosophy*.

Gutman and Schoon 7). These are not the only articulations of noncognitive traits, but arguably the most used and influential in the research to date. For the purposes of articulating goals for labor-based grading contract ecologies, I argue that there may be at least one more noncognitive competency useful to understanding learning and development: *a willingness to labor*. This noncognitive, as I define it, draws on three existing noncognitive traits studied already, creating a new noncognitive. In my courses, if students can demonstrate a willingness to labor, then my labor-based grading contract ecology has been effective.

The three noncognitive traits that a willingness to labor draws on are engagement, coping or resilience, and metacognitive strategies. And so, a willingness to labor, as a noncognitive competency, can be defined as a process of engagement that may start as simply satisfactorily doing the labor or work required in a course, but becomes a process of mindful, reflective labor that includes a desire or readiness to do that labor satisfactorily, despite setbacks and difficulties. In the past, I've called this a willingness to labor, "laboring to learn" instead of laboring to earn a grade (*Antiracist* 193-94). Here I attempt to theorize a willingness to labor through the research on noncognitive competencies.

Thus the overarching goal of labor-based grading contract ecologies, for me, is to get students to practice a network of interlocking, noncognitive competencies (engagement, coping and resilience, and metacognition), which I think of as a *willingness to labor*—that is, if all goes well, the particular goals that tend to regulate my labor-base grading contract ecologies encourage students to practice a willingness to labor, which are articulated primarily in Goals 1-3 below. This means that in order for me to understand how effective or successful my assessment ecology is, I must collect and consider evidence of a willingness to labor first and foremost. The design of my ecology, then, incorporates ways to do this, which have equal if not more benefits for my students.

GOAL 1: TO ENGAGE IN CONSISTENT, MINDFUL, AND MEANINGFUL PRACTICES

Gutman and Schoon place engagement under the umbrella term of "perseverance," and distinguish it from "grit" (17). They explain:

Engagement involves how students behave, feel, and think regarding their commitment to academic tasks, activities, or school more generally (Fredricks et al., 2004), while grit refers to a trait-level perseverance and passion for long-term goals which is related to Conscientiousness (Duckworth et al., 2007).

A willingness to labor, then, is not just an intention or knowledge set. Since it involves engagement, it is also action, laboring that affects behavior and behavior that affects laboring. This dialectical relationship between the intention of the student and their labor provides for beneficial behaviors, feelings, actions, and commitments. On one hand, it is a stance that affects the way a student approaches any labor asked of them in a course, allowing them to at least do the labor, even if they don't care for it, or find it interesting at first. On the other hand, it is laboring itself, which over time, changes the way the student thinks and feels about that labor, helping them find meaning and significance in it. Fredricks, Blumenfeld, and Paris explain that most conceptions of engagement as a noncognitive trait have three forms, or three kinds of engagement by students: (1) behavioral, such as following rules, asking questions, paying attention, and participation in school-related activities (62); (2) emotional, which "refers to students' affective reactions in the classroom, including interest, boredom, happiness, sadness, and anxiety," as well as identification with the classroom, peers, and/or teacher (63); and (3) cognitive, which tends to refer to a student's investment in learning, "self-regulation, or being strategic," for instance, "a desire to go beyond the requirements, and a preference for challenge" (63).

Thus asking students how they feel about a particular assignment or session of labor is vital to helping students notice and perhaps control or shape their engagement in a course. Early in each quarter or semester, I often pose the question: Why do you find particular activities or topics of interest more engaging than others? Why do you have the interests you do? What I want students to grapple with is how our interests and engagement in particular topics, for instance, have a dialectical relationship with the practices and activities that are associated with those topics. It is a chicken-and-egg question. Students might say they are interested in football or softball because they played these sports in high school. So I ask, why did you play them in high school? Did your playing create interest? How can you know you are interested in football or softball before you begin playing it? Noticing such feelings and attitudes toward the work of a course, during or after one does it, exercises metacognitive strategies, which I'll say more about below, but it also helps students notice when and how they are becoming engaged.

And engagement can look and feel different at various times in the course, for various activities, and to different students. Engagement is not always happy and exciting feelings about an activity or topic, nor is it always a sense of losing track of time in an activity. It can be boredom or fatigue, struggle or tension and anxiety. Engagement should be understood in nuanced ways and discussed with students as a full range of emotions, feelings, thoughts, behaviors, and actions. Much of our weekly labor and mindfulness journal entries and discus-

sions revolve around describing, naming, and talking through the ways we were engaged and how those feelings, thoughts, and actions help us or change the way we do our labors and our willingness to labor in the future for the course. The first two dimensions of three-dimensional labor in Chapter 3 (the *how* and the *that* dimensions of labor) provide reflective ways (prompting) to help students habitually access their ways of engagement in the course.

The most important thing in labor-based grading contract ecologies that I've cultivated over the years is not that students produce some ideal text or Discourse, or perform in some way that matches a preferred *habitus*, but that they engage mindfully and meaningfully in the practices of producing texts, which includes circulating those texts among readers, acting as readers themselves, and processing feedback from readers. This first goal, then, boils down to laboring at reading and writing practices in self-conscious ways. Asking students to labor in particular ways for certain amounts of time and simultaneously paying attention to that labor are the central actions of this goal. The assumption is that with the right amount of time and labor, any student will learn as much as they can. But it is possible that one can go through the motions and get very little out of a practice, so this goal also asks for mindful labor.

Being mindful and understanding one's labors as meaningful—i.e., full of meaning—are interconnected qualities of labors and often take time to cultivate. When a student sees their labor as meaningful, such as writing an essay for a class, they are aware—they notice things that have meaning, significance, or usefulness to them. Meaning is developed through self-conscious and reflective stances that they have engaged in during their writing labors or thought about afterwards. As I discuss in Chapter 3 as three-dimensional labor and mindful labor, this often means the labor of writing is done in a self-conscious way—that is, done while simultaneously realizing *that* the student is laboring at something in particular ways, at a particular time, under particular circumstances. Noticing such things in or after the writing labor accumulates meaning and value for the student, but it is very difficult to quantify how much meaning is accumulated or what it looks like (its nature). Reflective writing can offer a window into a student's mindful practices, but it needs to be done frequently throughout the semester, since consistency and repetition are important to cultivating mindful *habitus*.

What evidence in the course might there be of achieving this goal? The most direct evidence may be to measure the amount and/or consistency of labor done, then cross-reference that labor with some measure of mindfulness and meaningfulness. There are a few ways that seem most direct to me, but each have their problems. Chief among these problems is that most kinds of data on student labor are student-reported, and could be considered too subjective or inconsis-

tent to be useful for determining an ecology's effectiveness. Students can fudge self-reported numbers, even unintentionally, but a teacher can still use them to determine the ecology's effectiveness.⁵¹ The important thing is that you have to decide if you trust your students. If you do, then self-reporting information, such as labor logs and labor journals, will be enough evidence to make arguments about your ecology's effectiveness along this goal, but if you are using such data to make arguments to other stakeholders outside your classroom, then you'll have to decide what those audiences feel are strong evidence and/or how to frame any evidence used.

As one quick and surely incomplete example, consider a recent, writing course of mine. There were fourteen students enrolled in this first-year writing course, all first-year students in their first quarter of college. Most of the class worked outside of school. As a way to compare, consider the average combined labor accumulated by the bottom three (Latino, two White females), the middle three (Latino, Black female, White male), and the top three students (Latina, White female, White male). All students were between the ages of 18–20 years old.

Table 7.1. Average combined labor accumulated in a recent writing course by the top three, middle three, and bottom three students

	Total Labor (hrs)	Avg. Weekly Labor (hrs)	Avg. Reading Engagement	Avg. Writing Engagement	Overall Engagement	Avg. Final Course Grade
Top Three	82.03	8.20	3.23	3.89	3.56	4.0 (A)
Middle Three	69.77	6.98	2.98	3.45	3.21	3.7 (A-)
Bottom Three	42.28	4.22	3.23	3.03	3.13	2.6 (B-)

Let me emphasize that I am only identifying the amount of labor accumulated in the course when I identify them as top, middle, and bottom, as can be seen in the total and average weekly labor figures in Table 7.1 (first two columns). Those numbers determine bottom, middle, and top performers by amount

⁵¹ This is, of course, not the same as using these same measures to determine whether students have met a labor-based grading contract's obligations—as discussed in Chapter 4, I do not suggest using any self-reporting of labor (e.g., their labor logs) by students as a way to determine meeting contract obligations and thus determining grades.

of labor recorded in their labor logs. All students represented above met the minimum labor requirements of the contract and passed the course. The engagement ratings used a simple scale of 1–5, with 5 as most engaged and 1 as not engaged at all that students provided for each session of labor recorded in their labor logs.

What I find most helpful in determining consistent engagement in labor are the average and total labor hours accumulated by students. As mentioned before, the more a student labors in my contract ecologies, the higher the final course grade a student achieves, so the top three receive higher grades than the middle, and the middle higher than the bottom (the default grade for meeting our contract was a B, or 3.1).⁵² In the five unit course represented above, I assigned students an average of eight to ten hours of labor outside of class each week of the ten-week quarter.⁵³ While I did not use labor-logs to determine course grades, the average weekly labor in their logs shows that the top performers were able to meet these labor targets. And since engagement in writing labors are almost always higher than reading labors, I often use writing engagement ratings as one initial measure of engagement in the class (keeping in mind that this does not mean that a student had to like all labor to be deeply engaged in it), but even taking an overall engagement rating in this case shows an ascending set of ratings, with the top performers recording a higher average engagement rating. This is a typical distribution of grades and labor in my first-year writing courses. As the table shows, the more labor a student did, the more engaged they said they were, and our contract rewarded them with higher course grades—keeping in mind that I do not use labor logs to determine whether students meet our contract.

And how effective along this goal is my assessment ecology at helping my students of color? Is it socially just? How do students of color do? When I separate the data from Table 7.1 by race of student, students of color do more labor, and are engaged at the same rates as their White peers. The average total labor recorded for the quarter for the five White students represented in Table 7.1 was 56.71 hours, while the same average for the four students of color was 74.67. That's an average of almost 2 hours more labor each week that my students of color did. This places the students of color between the middle and top groups in Table 7.1, and the White students between the middle and bottom group.

52 The University of Washington uses a grading system that only allows a teacher to input numerical course grades of 0.0 or 0.7–4.0. You can see the scale on the university's website (https://www.washington.edu/students/genclat/front/Grading_Sys.html).

53 As one more reference point, in a more recent FYW course, one conducted after this one, the average labor per week of all students in that course was 9.40 hours and the average overall engagement rating was 3.10.

Keep in mind that most of the students of color in this sample were in the middle and bottom groups.⁵⁴

And what about average overall engagement? The White students recorded an average of 3.43 overall engagement and the students of color a 3.40. No difference. In terms of labor hours recorded, these numbers suggest that the ecology was effective at getting students to engage in the amount of labor I'd hoped for all students (eight to ten hours per week), with my students of color seeming to do more than my White students. In this way, I see my labor-based assessment ecology helping my students of color achieve as much or more labor than my White students, with equal engagement. Keep in mind the discouraging and unrewarding histories my students of color likely had in their past English and writing classrooms.

As I mentioned above, mindful, consistent labor makes for more meaningful labor, so seeing how much labor students accomplish in a semester or quarter is not enough to really understand how effective an ecology has been at achieving this goal. For now, I wish to point out that the mere act of paying attention to when, how long, where, and how engaged one is in their labors of the course through keeping a labor log and using frequent reflections on one's labor (see the section, "Three-Dimensional Articulation of Labor" in Chapter 3) is one way to see consistent, mindful laboring. So the fact that (1) all these students kept labor logs and labor journals consistently during the ten-week quarter; that (2) generally speaking, the more labor a student recorded the higher they rated their engagement in their labor; and (3) that students of color out-labored their White peers; all suggest that the labors were both consistent and mindful, and perhaps more so for my students of color. In the third goal below, I'll offer evidence from these students' labor journals and reflective writing in the course, but here, I'm suggesting that a teacher might measure the effectiveness of this first goal by looking at the data in labor logs and the presence of frequent, student reflections in the course.

But in order to confirm that any of the labor above is more than going through the motions, students must somehow engage with it in mindful ways, assessing what they were doing and how they were feeling. In their weekly labor and mindfulness journals that we write for just five minutes each week, I ask them to reflect upon one session of labor and explain how they were engaged with it. One of the students in the top performing group, T1 (White female), illustrates a typical kind of modest change in mindful laboring. A few weeks into the quarter, she writes:

54 I should also note that most of the White students in this sample completed their labor logs up to week 9, so their lower numbers could be accounted by this fact. However, as I'll discuss below in Goal 2, this trend may be just as important in understanding effectiveness of the assessment ecology.

the assignment I was most productive on was my precis. I really focused on it this week and tried hard to make it sound good, while still following the outline and having only five sentences. I wrote this in my room while I was in bed listening to some soft music. The session was engaging for me because I was in a time crunch with trying to get my assignment done before work, so I really had to focus on it and get it done so I wouldn't have to come home and finish it before I went to sleep.⁵⁵

T1 focuses on the time crunch that is created because of her work schedule. She associates engagement in the labor with how much time she has to complete that labor before she heads off to work. Her engagement is product-centered, dictated by the goal of finishing the job. Just a few weeks later, she reflects on a similar labor session:

I worked hard on meeting the requirement for the drafts and trying to give everyone in my group a good and thorough review of their papers. I did this by reading their drafts carefully and explaining how their papers made me feel and by focusing on the two aspect[s] of the rubric that they wanted me to focus on. I needed to take breaks between everyone's papers so I wouldn't lose my mind and become uninterested, so I guess by doing this it took me longer to get this assignment done. I also was distracted by my phone with texts and notifications I was getting, so I think while I do assignments I should probably turn my phone off or put it on Do Not Disturb so I can focus on what I am supposed to be doing a little easier.

Now, T1 becomes more strategic in her laboring. She notices when she's losing focus when reading her colleagues' drafts, takes a break, and returns to the labor with new focus. What now seems to construct her engagement in her labors is her focus, her sense of being interested in the work and her colleagues' papers. She recognizes when she is interested and even finds a new practice to help her, turning her cell phone on "Do Not Disturb" so she is not interrupted. This new articulation of her labor is centered on not a product, like getting the assignment done, but on doing the labor in a focused and interested way.

While I do not wish to make too much of these numbers, nor suggest some correlation between the change in laboring habits that T1 shows in her reflec-

55 I do not edit any student writing offered in this chapter, except on a few, rare occasions for readability purposes.

tions, I do wish to point out that all three groups found the writing labor of the class engaging to some degree. All the students in each of the three groups above offered similar kinds of reflections, with a gradual shifting in their reflections over the quarter in the directions that T1 illustrates. They generally moved away from product-oriented, get-the-job-done goals for their labor sessions, and toward finding interest and focus, and noticing particular feelings arise during labor sessions (not always good or happy feelings).

Many of these changes may have been nudged by my prompting in class, which emphasizes describing a labor session, reflecting on how engagement was experienced, felt, and created in the session of labor, then finding something that the session teaches them for their future labor practices (discussed in Chapter 3). This prompting tends to call our attentions to several aspects of engagement in labors: (1) that engagement is not always happy or easy, but can also be feelings of difficulty and struggle; (2) that engagement can be created by consciously arranging one's body, laboring environment, or circumstances; and (3) that engagement is not a random occurrence that surprises us (it doesn't just happen to us), but something we might plan or construct in a premeditated fashion.

Finally, I've found more recently that engagement is a good initial topic of discussion to have with students, particularly since I'm asking them to rate their engagement on a scale of 1–5. In the past, I've simply given them the scale and a general guideline for each rating, such as: 1 = very low; 2 = low; 3 = neutral; 4 = high; and 5 = very high. But this scale doesn't account for the nuance that I'm describing above, nor the nuance in reflections and journal entries. So opening a discussion early on to define what each student might observe or account for in their labor sessions, then make note of what those feelings and experiences are somewhere, and frequently returning to the definition of engagement may offer students more nuanced, and perhaps more internally consistent, ways of recording and reflecting up engagement.

Table 7.2 offers various kinds of evidence that I find useful in determining a labor-based grading contract ecology's effectiveness along this first goal, which might be gathered at the end of a semester or quarter, as well as examined with students at the midpoint of the term or semester. As Chapter 3 explains, such data can be useful for students themselves to reflect upon at strategic points in the semester or quarter. Table 7.2 links particular prompts or questions to the data in labor logs. The reflections students produce might do a number of things beyond provide the teacher with student-generated evidence for the degree of effectiveness of the ecology. For instance, journal reflections can help students analyze or make sense of the numbers, looking empirically at their engagement through total labor time in the class, average engagement ratings, or how those numbers change during the course of the semester or quarter.

Table 7.2. Kinds of evidence for engagement in consistent, mindful, and meaningful practices

Data from Labor Log	Questions About Data
Amount of minutes/hours of labor accumulated in the course	How much labor do you ask of/assign students each week? How close do students come to this target? How much labor has been reading sessions and how much writing sessions? Are there differences among various identifiable groups in the class (e.g., race, gender, generation of student)?
Average minutes/hours of labor per week of the semester/quarter	How much labor does each student average each week? What differences in averages are there among identifiable groups in the class? How consistent are the weekly averages across the quarter or semester? Are there differences between identifiable groups? How do you account for these differences?
Average engagement ratings of labor sessions	What are the average reading, writing, and overall engagement ratings for all students? What differences do you see between identifiable groups? Do overall or group average engagement ratings stay consistent across the quarter or semester? How do you account for any differences?
Comparisons of amount of labor to engagement ratings	Considering just the labor sessions that students rated highest (5) and lowest (1 or 2) in engagement, what is the average amount of total labor (or average amount of time per session) in identifiable groups? How does the amount of time compare to the mean (average) or the median (student who represents the exact middle value) of the entire class?

GOAL 2: TO CONSCIOUSLY LABOR AND WORK TOWARD RESILIENCE

This second goal of my labor-based grading contract ecologies centers on the noncognitive competencies of coping and resilience. Coping and resilience have been used interchangeably in the noncognitive literature, however, they are distinct concepts, “reflect[ing] distinct aspects of successful development and adaptation” (Compas et al. 89). Gutman and Schoon explain, “resilience and coping are both concerned with how individuals respond to stress, [but] they are conceptually distinct. Coping involves skills that people use when faced with specific difficulties, whereas resilience is a process which follows the exercise of those skills” (27). Lazarus and Folkman explain coping as “constantly changing cognitive and behavioral efforts to manage specific external and/or internal

demands that are appraised as taxing or exceeding the resources of the person” (141). When one successfully copes with stressful or difficult situations, one becomes more resilient, willing to continue to exercise coping strategies in the future. Resilience, then, is a process of “positive adaptation despite the presence of risk” (Gutman and Schoon 27) through coping behaviors. I think of coping as the root noncognitive competency in coping and resilience. Thus, coping strategies and behaviors lead to resilience, or positive adaptation to stressors, risk, and difficulties. Often coping in my class is simply laboring in ways we’ve asked of each other.

While coping is the root noncognitive competency and resiliency is the process of successful coping, I see these two noncognitives together. For my ecologies, coping and resilience are a linked process of “purposeful responses” that move toward a sticking to tasks and labors and eventual success in tasks or goals in some way (Gutman and Schoon 27) that is recognizable to some degree by the student. Compas et al. offer this definition of coping, which I think is useful for understanding coping and resilience as a single process in a willingness to labor:

conscious volitional efforts to regulate emotion, cognition, behavior, physiology, and the environment in response to stressful events or circumstances. These regulatory processes both draw on and are constrained by the biological, cognitive, social and emotional development of the individual. An individual’s developmental level both contributes to the resources that are available for coping and limits the types of coping responses the individual can enact. (89)

Thus everyone enters a given writing course with different experiences and coping competencies, all of which are built upon during the present course. These coping competencies allow the person to regulate their emotions, cognitive abilities, behaviors, their psychology, and the environment around them. In their meta-analysis of 124 articles that studied coping, 165 independent samples, which included 33,094 participants in studies between 1980 and 2004, Connor Smith and Flachsbart identified two broad forms of coping, “approach” or engagement, and “avoidant” or disengagement (Connor Smith and Flachsbart 1081). Coping, then, appears to come in at least two forms, engaging with stressors in some way or avoiding them. The literature is clear that avoidance coping strategies have detrimental or negative effects on people and their abilities to become resilient (Barlow, Allen, and Choate 219-21), while approach or engagement coping has positive effects, helping people become more resilient through difficulties.

In their study on positive effects due to approach coping strategies after participants went through mindfulness-based interventions, Cousin and Crane explain these two kinds of coping strategies:

Engagement strategies include responses that are oriented towards the stressor or one's reactions to it, such as problem solving, seeking for social support, or acceptance (i.e., coming to live with the stressor). Disengagement strategies, on the contrary, include responses that are oriented away from the stressor or one's reactions to it, such as avoidance (i.e., trying to avoid facing the problem, or the thoughts or emotions related to it) or denial. By encouraging awareness of one's experience in its entirety and acceptance of thoughts and affective states whatever their form may be (Segal, Williams, & Teasdale, 2012 [*sic*]), mindfulness meditation prevents disengagement from difficult experiences when they arise (435-36)

While I'll say more about mindfulness and metacognitive strategies in Goal 3, it is important to note the close association in the psychological literature on coping. Metacognition and awareness, in the above case through meditation and mindfulness practices, appear to offer approach coping strategies. In other words, people who went through an eight-week mindfulness-based intervention (a course) gained approach coping strategies, and reduced avoidant ones (Cousins and Crane 443). These changes in coping strategies were statistically significant, and resulted in positive effects on participants. In part because they practiced noticing their responses to stress and difficulty and their coping behaviors, participants found more effective coping strategies.

Ultimately, approach coping, if it is engaged in with some degree of conscious awareness, is an integral part to a willingness to labor. It explains a stance or orientation that a student moves toward in the labor-based assessment ecology, one that does not avoid labor or difficulties, but one that becomes oriented toward difficult or challenging labor and problems, in part through awareness. My assumption is, and I could be wrong, that many students, particularly those most harmed by past writing assessment ecologies, have a higher frequency of using avoidant coping strategies, like not doing enough work in a course, starting late or procrastinating on assignments, or not engaging fully in class activities, reading, or writing assignments.

Evidence of the effectiveness of the ecology along this goal can be seen in similar ways as the first goal. Students engage in approach coping practices, first, by simply doing the labor asked of them without the pressure of grades of quality on their writing. As mentioned in Chapters 3 and 4, the labor instructions I

provide for each reading or writing assignment reminds students in its opening description, purpose, and goals statements that what I'm mostly asking them to do is labor in particular ways, to be in the labor, not to finish something and turn it in, even though they will do that (see Appendix D for example labor instructions). I reinforce coping strategies by asking students to pause for a minute in their labor processes, assess what they've just done, how they've done it, and what they may be feeling, then Slack/tweet that to the class. Often, simply noticing what one has accomplished can be a coping practice, a tacit validation that shows the student, "yes, I am doing this, and I can do this work."

Gradually over the course of the term, I shape the prompts for these labor tweets to help students try out at least two other approach coping practices. First, I ask them to reframe the current work (if it is difficult work). *Reframing* could be altering how one thinks about the current work, rearticulating the goals one sets for the labor, or consciously changing one's values that are linked to the labor at hand. For instance, when reading a difficult or dense text, one with a lot of theory or new terms and ideas in it, I may ask students to pause at some difficult point in the reading (e.g., "after reading page 39, pause and tweet . . ."), spend two or three minutes to reframe their goals for reading the next few pages of the text. If we started our reading with the goal of understanding to summarize the text, I might suggest at this point that a student now read the next few pages in order to tweet a short poem or haiku using at least two or three words/terms that are new to them from those pages.

The second coping practice I offer in labor Slacks/tweets is to simply notice and accept their current feelings about the present work in a non-judgmental way (e.g., "I'm bored," or "this is confusing to read"), affirming that their feelings about their labor does not mean they cannot do the work or are a bad student, reader, or writer. We might call this coping practice *acknowledging feelings*. It is one way to remain embodied as we do otherwise intellectual work. The tweet might be taking a picture of a page of the text, and telling us how they felt while they read that page. The point I try to help them see is that their feelings do not dictate their success in laboring, in coping and being resilient. The labor tweets and reflective moments around our labor, such as the labor logs, help us further notice how we are coping, and through noticing, students can build more approach coping competence, which leads to resilience. Because approach coping practices are so vital to laboring toward resilience, thus to the larger noncognitive of a *willingness to labor*, I reiterate these four teaching practices:

- **Focus assignment instructions on the process of labor** to accomplish and the amount of time on task, words to write or read, and not on some standard of quality

- **Integrate mindful pauses** in labor processes to help students notice what they have done, how they did it, and what they are learning.
- **Ask students to reframe the labor** or work at hand by rearticulating it in a different genre, to a different audience, or from a different perspective from the ones they started with.
- **Integrate ways to acknowledge students' feelings** about the labor at hand, allowing students to voice in some way their emotional responses to the work of the course.

The assumption is that most students, once they stop getting grades and other external motivators to guide their labor, they must replace those motivations with others. Approach coping practices are the ones I focus on. The previous examples of journal entries by T1 illustrate a more qualitative way to see evidence of coping in the ecology. T1 shows how she's not only able to stay in the labor by noticing how she is coping with the challenges of getting her labor done (e.g., she takes breaks to maintain focus in her reading), but also how she invents new ways to cope (i.e., to turn her cell phone to "Do Not Disturb").

M1, a Latina student from the middle group, reflects in her labor and mindfulness journal around the same time in the quarter (around week 6–7), showing coping in subtle but similar ways. M1 reflects:

One labor session that i felt the most interesting and engaging for me was the extra three precis, even though they weren't required to do i felt like doing 5 wasn't really enough for me to understand our texts and what we read. While writing these i was in my bed chillin with criminal minds in the background as usual, but unlike our other precis these ones seemed easier to do. I didn't have to pause or mute my show like i usually do to focus on these and instead writing them just seemed to flow this time around. I feel like it was more productive because it wasn't really required. I didn't start thinking about "well do i really have to do these?" instead i found myself more engaged and interested also because I didn't have a specific reading i had to write about. I intend to use this motivation during the precis with my dramatic revision so that my writing can flow like they did with this labor.

M1's reflection appears at its face to show less explicit coping with the labor of the class. Instead of consciously adjusting her environment or behaviors to try to change her engagement in the writing of her precis, she relies on whatever feelings arise. In this case, they happen to help her because these precis were

extra labor (done to get a higher course grade). M1 just notices and hopes she can keep up this kind of engaged labor. She notices what helps her focus or stay engaged in the activity, but has a vaguer sense than T1 of how to control her future labor sessions, or does not discuss how she might. While not as explicit as T1, M1’s reflection still shows coping through noticing her engagement, and acknowledging her feelings as she labors. She was chillin in her bed, engaged and interested in the work, and she notices that she’s not thinking, “well, do I need to do these?” She acknowledges this as a feeling (“I feel like . . .”).

Furthermore, when I look at M1’s actual labor accumulated by week in the quarter next to T1’s actual labor, and B1 (a Latino student in the bottom performing group), some interesting patterns appear that suggest consistent coping and resilience behaviors by M1.

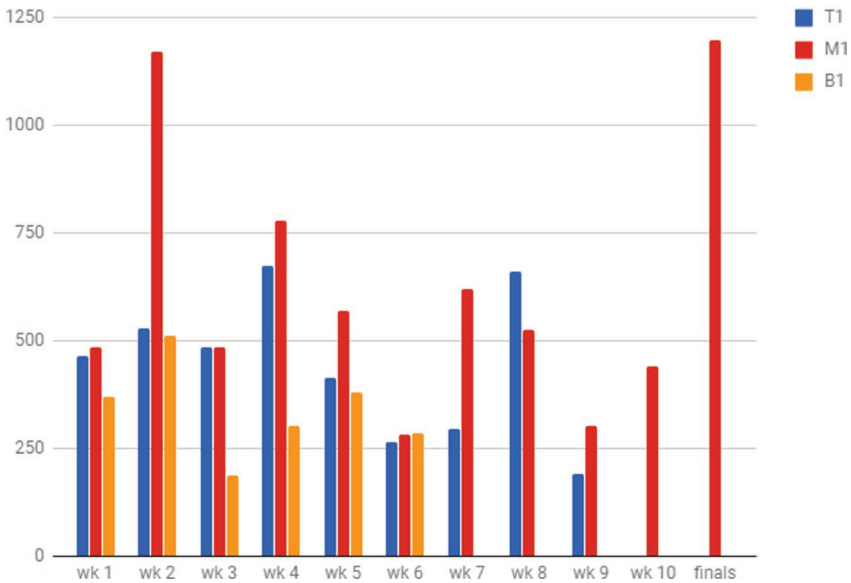


Figure 7.1. Actual recorded labor of T1, M1, and B1 by week of the quarter.

One coping strategy I’ve been suggesting is the simple act of noticing one’s behavior, then recording it in labor logs (that’s a signal of noticing). T1 didn’t record her labor for the final week or during finals week, and B1 neglected to record his labor from weeks 7 on. Meanwhile, M1 records her labor through finals week. One way to read these data is to see the effects of M1 noticing her labor. They help her cope and be resilient to the end of the quarter. One confirmation of this is the amount of labor she records. Her labor did not decrease at the end of the quarter. In fact, it ramped up. All three students had similar labor patterns during the quarter, but M1 consistently recorded more labor than her peers.

She was able to stay in the labors, which allowed her to do more, and continue recording her labor.

And based on the higher amount of labor recorded by the students of color in this sample, M1 (Latina) is not an anomaly when it comes to coping and resilience. M2 (Black female) shows a similar pattern in her labor by week. She recorded 8.5 hours, 7.67 hours, and 16 hours of labor in the final three weeks of the quarter (weeks 9–10 and finals), a ramping up of labor. Meanwhile T3 (Latina) recorded 8.08 hours, 7.75 hours, and 6.67 hours in the same weeks of the quarter. While not a ramping up, it is a consistent recording to the end of the quarter. The only White student who completed his labor log to the end of the quarter was in the top group, T2, and he recorded 15 hours, 10 hours, and 5.17 hours in the final weeks—a ramping down. The same trend that T1 seems to be headed toward in Figure 7.1. All the other White students in the sample stopped recording their labor sessions in about week 9, where T1 stops. While this does affect the labor totals, it shows a trend. Is it possible that the act of completing the labor logs to the very end is a coping strategy that works well for students of color?

I want to emphasize that the actual numbers are less important in labor logs, since I don't think it is wise to compare exact numbers between individual students. Averages of identifiable groups (and the whole class) and general patterns (trends), such as those easily seen in Figure 7.1, are more useful in considering effectiveness of the ecology. Maintaining and ramping up in labor suggests successful coping and resilience strategies. Not being too focused on the numbers keeps a teacher from using one group's averages as a criterion by which all other groups in the class are judged effective or not. However, I do have a target number of hours per week I shoot for since I have identify estimated labor time on all labor instructions, so I do use this as a general guideline to understand whether I'm asking too much for a particular class or not enough labor during the quarter. Knowing how much on average you were asking students to labor each week during a term can be a useful criterion by which to gauge how effective the ecology was along Goal 1 and provides a guideline for reading labor log data for evidence of coping and resilience (Goal 2).

Many of the students in this class worked, and some had family obligations, while others only went to school, having more time to spend on school work. These factors changed what students can record, and may have actually enhance their coping competencies, since they needed more useful ones to succeed in school under heavier schedule obligations. During our discussions of our labor log entries in the middle of the quarter, around week 6, we changed the number of estimated labor each week from around twelve to around nine hours. Regardless of the actual numbers for any given class, a teacher might expect that the

more students labor, generally speaking, the higher the likelihood that there are successful approach coping practices present in that group of students. The more labor recorded, and the longer they keep their labor logs also suggest coping and resilience practices.

Table 7.3 offers various kinds of evidence that may help understand how effective an ecology has been along this goal.

Table 7.3. Kinds of evidence of coping and resilience

Evidence of Coping and Resilience	Questions About Data
Amount of minutes/hours of labor accumulated in the course	How many students in the class were able to come close to or exceed the labor goals (in total mins/hours) that were agreed upon for the quarter/semester? What do students say about why they did or could not meet those labor targets? What approach coping practices did students try and say were helpful to them in accomplishing more labor? Do students complete their labor logs to the very end of the term?
Average minutes/hours of labor per week of the semester/quarter	What was the average total labor (mins/hours) for each student per week? How many students consistently achieved the labor goals (in hours/mins) each week? What do students say in weekly labor journals causes their labor time to dip below the minimum expectations? What coping strategies did students employ to maintain consistent labor week by week?
Average engagement ratings of labor sessions	How many students achieve 3 or better average engagement ratings? What do students say in journals that those ratings mean each week? Are their engagement ratings consistent over time or change? How well were students able to separate their feelings about labor from their sense of being successful in accomplishing the labor required of them?
Students' reflections on coping strategies	When asked periodically, what do students say about how they cope with the labor expected of them in the course? How closely do they say they follow the labor instructions given for assignments? Why do they deviate from them? Do they describe deviations in labor practices as approach or avoidance coping strategies?

Eli Review, a web-based, student, peer-review program that manages feedback and revision cycles for drafts, offers other kinds of data on the fly that tell stories about student intensity of work and engagement, much of which is determined by amount of words produced in feedback, number of comments made on peers' drafts, and the number of their comments that are judged effective or helpful by writers receiving feedback. If read in the right ways, this data too could tell something of coping and resilience. But as of the writing of this book, Eli Review only accounts for feedback processes, so it cannot say anything about coping or resilience, for instance, in any other labors of the course, like reading or even drafting and revising one's own drafts. Still, technologies that offer such data on the fly can be helpful in discerning such coping (Goal 2) or engagement (Goal 1) practices during a semester or quarter.

GOAL 3: TO PRACTICE METACOGNITIVE STRATEGIES FOR UNDERSTANDING ONE'S LABOR PRACTICES

While metacognitive competencies are a part of Goals 1 and 2, it is important for me to measure metacognition as a separate ecological goal. This means as a teacher, I attempt to gather evidence just to assess how effective the ecology is at encouraging students to practice metacognitive strategies as a way to understand their labor practices. In Chapter 3, I discussed labor as a three-dimensional, mindful practice, which maps most closely to the noncognitive literature on metacognitive strategies. Metacognitive strategies are those that involve a number of learning processes, such as goal-setting, planning, problem-solving, reflecting on one's strengths and weaknesses, monitoring progress, and making decisions about certain tasks and practices (Gutman and Schoon 22). Paul Pintrich provides a distinction between various definitions and models of metacognition in the psychological studies of it. There are those that focus on "knowledge of cognition" and those that focus on "the processes involving the monitoring, control, and regulation of cognition" (219).

In his overview of the literature around assessing metacognition in learning environments, drawing on Flavell's work on metacognition and cognitive monitoring, Pintrich offers a three-part model for metacognition. Metacognition, he says, involves, first, "strategic knowledge" about "general strategies for learning, thinking, and problem-solving" that tend to work across various domains and disciplines (220). For example,

Strategic knowledge includes knowledge of the various strategies students might use to memorize material, to extract meaning from text, and to comprehend what they hear in

classrooms or what they read in books and other course materials. Although there are a large number of different learning strategies, they can be grouped into three general categories: rehearsal, elaboration, and organizational (Weinstein & Mayer, 1986). (Pintrich 220)

The second part to Pintrich's model is "knowledge about cognitive tasks," which include information about particular tasks' difficulty and other cognitive tasks required within a larger task. Knowledge about cognitive tasks is understanding the "what," "how," "when," and "why" of particular cognitive tasks (Paris et al. 296-98; qtd. in Pintrich 221). This knowledge helps a student use particular strategies and knowledge in particular situations for particular purposes. In writing classrooms, we often help students with rhetorical awareness and reading texts rhetorically, which are treated as cognitive abilities, but Pintrich's model explains that those cognitive abilities move into, or share, noncognitive domains, likely those cognitive competencies become noncognitive over time as a practice becomes "second nature" or habit (or as I've said here and elsewhere, as it becomes *habitus*). Finally, Pintrich identifies accurate "self-knowledge" about one's strengths and weaknesses as the third element in his model (Pentrich 221).

So if labor is to be mindful in the ways that I discuss in Chapter 3, then another way to explain the three dimensions of that mindful labor is to see those dimensions of labor as places to exercise metacognitive practices. The nature of these practices, what they activate or exercise in students, is at least three things:

1. strategic knowledge and knowledge about learning, thinking, and problem-solving;
2. knowledge about specific cognitive reading and writing tasks (the what, how, when, and why of tasks); and
3. accurate self-knowledge (strengths and weaknesses) about one's own reading and writing labors.

As you might guess, just keeping a labor log and journal offers evidence of each of these metacognitive dimensions, but it's the nature of that evidence and what students do that often matters most in understanding how effective a learning environment is along this kind of dimension. Having accurate self-knowledge about one's reading and writing labors (or understanding clearly the way others judge or would judge one's writing, for instance) is one thing, but using accurate self-knowledge strategically for planning and problem-solving in the future is another. Guided reflection is often required to make such knowledge usable or transferrable.

And so, reflection and metacognitive practices tend to make the habitual, our unseen or unheard practices and dispositions, seen and heard. Doing so gives a writer more control over their dispositions, providing knowledge of how to use them strategically, and perhaps change some of them. One way to think of this movement from cognitive to noncognitive is to think of it as a metacognitive process that pays attention to our labors and feelings during and after our labors. In labor-based grading ecologies, metacognitive practices often focus on the processes of labor and the attending attitudes and feelings experienced with them in order to build approach coping strategies and engagement, which allow students to stay in labor longer and have more opportunities to understand their labor in all three dimensions.

Reflections on one's labor, both as one is laboring and afterwards, then become evidence of metacognition. In fact, the act of recording one's labor is an act of metacognition, particularly in a labor log that requires the student to notice various features that construct their labor sessions, such as duration of the session, date and time of day, description of the labor, where the session of labor takes place, and engagement ratings. Pausing to record such data after each session is a metacognitive practice, just as labor tweets in the middle of a labor process are. Thus evidence of metacognition is not just written reflections, although those are important, but the data recorded by the student on their sessions of labor for the class.

Additionally, there has been much discussion in writing studies about the usefulness of metacognition for learning to write. In composition studies, the literature on reflection, and most notably Kathleen Blake Yancey's work (*Reflection; A Rhetoric*), offer writing teachers ways to think about metacognition in literacy classrooms. Yancey explains that in the different literatures on metacognition and reflection, these two concepts are defined differently as learner practices. Yancey says that metacognition is "thinking about thinking associated with planning, self-monitoring, and self-regulation," while reflection is oriented to self-assessment activit[ies] occurring at the end of a learning cycle, though capable of prompting a new one" (*A Rhetoric* 6). In order to practice metacognition, she explains that most models say that students should engage in processes of monitoring and control, thus they usually take on a pattern like this:

- assess the task at hand, taking into consideration the task's goals and constraints;
- evaluate their own knowledge and skills, identifying strengths and weaknesses;
- plan their approach in a way that accounts for the current situation;

- apply various strategies to enact their plan, monitoring their progress along the way;
- reflect on the degree to which their current approach is working so that they can adjust and restart the cycle as needed. (Yancey, *A Rhetoric* 7; Ambrose et al. 191-99)

Thus, in the psychological and educational literature, metacognition is a process of monitoring and controlling how one goes about doing the labors of learning, while reflection is the final step in the larger processes of metacognition. The kind of reflection that writing teachers tend to think of as metacognition, then, is just the last step in the above five-step process that is always happening, if a student is exercising metacognition along the way. To be clear, for my purposes in this chapter, metacognition occurs *in the labors of learning*, and is exemplified in the labor log, while reflection happens *after the labors of learning* but focuses its attention on those labors, and is exemplified in the labor journal and other formal reflection documents.

Of course, this is not the only way in which reflection has been defined. The two most influential figures for Yancey's work on reflection are John Dewey and Donald Schon. Dewey says that "reflective thought" is based on "active, persistent, and careful consideration" (qtd. in Yancey, *A Rhetoric* 7; Dewey 6), so it could be argued that Dewey is thinking about reflective thought as a part of a larger process of metacognition, perhaps similar to the one identified in the educational literature by Yancey. Schon's work on reflection among practitioners in professional settings explains reflection in at least two ways, as "reflection in action" (Schon 28) and as "reflective transfer" (Schon 97). Again, Schon also appears to be thinking about reflection as a part of an ongoing process that happens as one practices and that changes or "transfers" new ideas, practices, strategies, and competencies, which occurs afterwards. Yancey's own three kinds of reflections in writing classrooms is the most influential in writing circles and draws on Dewey and Schon. She defines three kinds of reflection for writing classrooms:

- *reflection-in-action*, the process of reviewing and projecting and revising, which takes place within a composing event;
- *constructive reflection*, the process of developing a cumulative, multi-selved, multi-voiced identity, which takes place between and among composing events; and
- *reflection-in-presentation*, the process of articulating the relationships between and among the multiple variables of writing and the writer in a specific context for a specific audience (Yancey, *Reflection* 200).

One might view Yancey's three kinds of reflection in writing courses as a loose overlay to the five-step process of metacognition mentioned above, with a more focused consideration on what happens to the learner in the reflection process. That is, Yancey's model focuses on the nature of the reflective discourses and practices that occur in the processes of learning to write. So her model actually is more labor-oriented, and reveals the nature of the changing reflective stances a student moves in and out of.

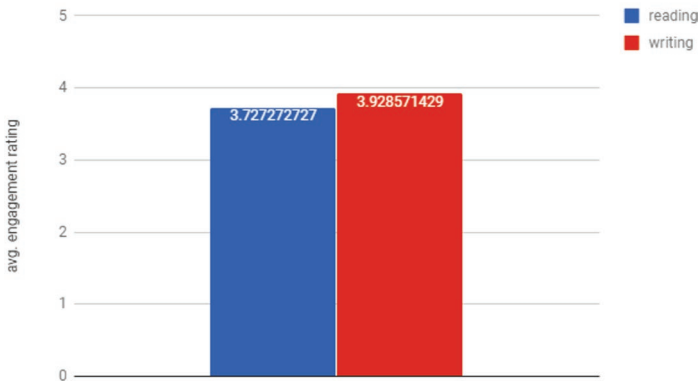
Yancey's three reflection practices also can be mapped to the metacognitive processes of learning that I've discussed already. Reflection-in-action is the metacognitive processes of monitoring and control that occur in the first four steps of Ambrose et al.'s model (above), cited by Yancey, and is equivalent to Schon's reflection in action. This metacognition, I attempt to capture in labor tweets that occur during the processes of labor—often, I think of them as modified versions of think aloud protocols, mindful moments that help us pay attention to our labors. Constructive reflection is the in-between process of metacognition, often happening along the way in a course periodically. I see this mostly in labor journal entries when students construct and articulate versions of themselves and their labor practices. While reflection-in-presentation is the final step in the course's movement to the end of the semester or quarter. In my courses, this often occurs formally in our end-of-quarter letter of reflection that pays careful attention to our labor as a practice, which examines labor logs and journals for patterns and insights, then represents labor as a practice over the semester or quarter by the student.

There are several ways to structure a course so that it is easy to assess an ecology's effectiveness along this course goal, which may be clear already. I simply make it difficult for me to not see evidence of metacognition, since we look consistently and carefully at our labor logs, labor tweets, and labor journals, then center our final conferences on their final, formal reflections on their labor as a practice during the quarter. These final reflections on labor as a practice, as discussed in Chapters 3 and 4, offer evidence of the ecology's ability to help students become more conscious of their labors in ways that help them monitor, plan, read, and write flexibly in the future.

The students in my course example from Table 7.1 generally had trouble with articulating their labor as a practice, which isn't unusual for young, first-year writing students. But a few students in the middle category (and most in the higher performing category) were able to begin doing this. For instance, M1 (Latina) offers this reflection during week 7 of the quarter. I asked students to look at their labor logs and make some observations about their labor as a practice. She writes:

Throughout these 7 weeks of the quarter I have produced 3,510 minutes worth of labor which divides into 510 minutes per week. I spend more time and I am more engaged on the work where I have to write rather than the reading. That makes me wonder why, is it because of the reading material? Or maybe I just learn better when I can write it out rather than when I read it? I liked looking at my labor and seeing how much I actually got done over the past 7 weeks. It makes me feel better knowing I actually did work and didn't slack off like I did when I did when I was in High School.

Average Engagement in Reading & Writing



Recording my labor has been very helpful in the way that I can see how much work into my assignments and see what I enjoyed doing most and what I didn't enjoy as much. What I did enjoy I obviously put more work into which means if I can find some way to be interested in whatever my assignment is than I can make it better and do better on the assignment. I plan to use that method with my other classes and assignments.

M1 does a good job here, well enough to find some useful information for her future practices. She demonstrates metacognitive practices by keeping her labor log current (so she can do this reflection) and by finding some useful patterns in her labors. In the process, she constructs in this journal entry a reflective self and description of her laboring in the course.

It is interesting to note that her average engagement ratings for reading and writing sessions are pretty close, only a difference of .2. More typically, students self-report their writing engagement at .5 or higher than their reading engage-

ment in my courses at UW Tacoma. And since both of these engagement ratings are well above 3, her labor sessions seemed to be going well. I consider engagement ratings of 3 or higher to be good engagement. What she isn't able to talk about, but hints at in her first paragraph, is that her reading labors total only about 1/7 of her total writing labors, meaning she engaged in seven times the amount of writing labor, than reading. Granted the course is set up to be more writing heavy, but not this lopsided. More typical in this class was for writing labors to be four or five times as much as reading labors. It likely meant for M1 that she found deeper engagement when she wrote for our class and so she did more of it. Her knowing this is a good sign of metacognition about her labors. And because M1 is in the middle category of students, one could argue that this is a good sign of the ecology's effectiveness along this goal. A deeper look at the reflections of others in this category (middle category) in the class would help understand how true this is.

But understanding how effective the ecology is at getting students to engage consistently in metacognitive practices about their labor can be measured in other ways too. Counting the number of labor Slacks/tweets each student accumulated is one way. For instance, the average number of labor tweets for the high, middle, and bottom performing groups from Table 7.1 align with their other labor data. The high group averaged 29.3 tweets over the semester, the middle group, 21.6, and the low group, 16.3. If a teacher wished to use labor tweets even more strategically toward metacognitive ends, making them a more integral part of the class and every assignment, other data could be gathered to understand how effective the ecology was at achieving this goal. For example, one possible index of metacognitive practice about one's labor could be total hours of labor divided by the total estimated hours assigned, then multiplied by the total number of tweets a student sent divided by the number assigned. Since an increase in either labor hours or tweets would increase the total index number, and more tweeting or labor continues to raise the score. One could weight either side of the equation further by using a multiplier, for instance, multiplying the tweets by two. As an example, here's my metacognitive index formula:

In my example groups from Table 7.1, the metacognitive indexes for each student in each category match my sense of each student's metacognitive capacities, as demonstrated in the course and their final portfolios. For example, T2 (White male) offered a total of 31 tweets in our 10 week quarter, and logged 92.67 hours of labor, which gives him a metacognitive index of 1.28. Meanwhile, T3 (Latina) had fewer labor hours logged (88.17) but substantially more tweets (37), which increased her index to 1.45. So while T3 didn't do quite as much labor as T1, which would be measured as favorably along the first goal above, along this third

goal, her labor tweets show her to be more metacognitive in the labor she did complete (she paused and tweeted more).⁵⁶

The metacognitive index is just a crude number that may help at midpoint, for instance, in seeing quickly how a class or individual students, are doing along this dimension, perhaps showing whether a teacher might need to probe more deeply into journals or have conversations in class. If used, take it with a grain of salt. The index is just a number, a construct. But I will say that when I average each group in Table 7.1, I get consistent metacognitive indexes that match generally my senses of those groups of students in my class (top is at 1.10, middle at .77, and bottom at .31). These indexes also suggest how close to the labor hours and labor tweet targets I tried to set for them, with a metacognitive index score of 1 being perfectly met.

Is there an index number that suggests to me a student isn't doing enough along this dimension? This is hard to say. Perhaps, but one would need some criteria by which to compare these numbers. My goals in this class were always to assign between eight to ten hours of labor each week, give or take, and to ask students to do on average 2.5 labor tweets each week. So if a student did ninety hours of labor and twenty-five tweets, they would get a metacognitive index of 1. But I'm often wanting students to tweet more, since I find that tweeting is a mindful activity when done in the way I ask. So I can see myself adjusting this formula to weigh more heavily tweets, say with a multiplier of two. This would mean that tweeting is calculated as more metacognitively heavy.

I offer these details about my metacognitive index, not so that you might use it yourself, but to show how a teacher might collect such simple data in order to measure such a goal empirically, even if it says very little about the nature of the metacognitive practices circulating in the ecology. This number, then could be evidence of an ecology's effectiveness along this goal, but a poor indicator of the nature of those metacognitive practices. Since I don't use outcomes, instead more open-ended goals, I am okay with only knowing that my ecology is effective in some way.

Finally, another kind of evidence a teacher might use to make arguments about their ecology's effectiveness along this goal might be simply the number of students who completed their labor logs to the end of the course. I used to have about 30%–50% of my students not complete their labor logs to the very end of

56 I should note that I've experimented with this formula for metacognitive index and found other useful formulas using the same data. For instance, a simpler version is one that replaces the right-side of the present formula with just number of tweets. So a theoretically typical student doing all that is asked of them, and nothing more, would have a metacognitive score of approximately equal to the number of tweets assigned at any given moment in the quarter or semester. The higher their score, the more metacognitively engaged they are.

the course, as was the case for this course.⁵⁷ Now, I do a better job at reminding students and using the logs each week, which has translated into nearly everyone completing their labor logs to the final days of the course. Table 7.4 shows several ways to gather evidence of effectiveness for this goal, and some questions that may help a teacher investigate that data.

Table 7.4. Kinds of possible evidence of metacognition around labor

Evidence of Metacognition	Questions about Data
Labor tweets	How many labor tweets per assignment (or per week) do students do on average? How many total did they do and how many did you ask of them?
Labor logs and metacognitive indexes	How many students completed their labor logs during the semester/quarter? How many entries per week did they do on average? What is the average metacognitive index for the top, middle, and bottom groups of students (determined by amount of labor recorded in labor logs)?
Labor journals	How many students completed their labor journals (did all journal entries assigned)? How many students had longer journal entries in the final few weeks of the semester/quarter than in the first few weeks?
Formal reflection documents	How many students were able to make sense of their labor logs and journals in formal reflection documents (journals and/or letters of reflection in portfolios)? How many were able to articulate future best practices or lessons learned about their reading or writing labors in those reflection documents?

Note that I'm not asking questions about the nature of metacognition in the evidence listed in Table 7.4, but one might if a teacher wanted to know the nature of the metacognition circulating in their course, which may be important to understanding the effectiveness of metacognition in that class (if metacognitive practices are predefined). As I discuss in Chapter 6, I don't know where each of my diverse students start in my course, nor can I anticipate what their linguistic competencies will be after my class. They learn what they can, and only that. I cannot force things, nor am I willing to punish students with lower grades for being who they are. This means that my ecology's effectiveness need only

⁵⁷ In this course, four of the twenty students did not make entries in their labor logs into November (finals week is early December). Interestingly, one of the students who completed the necessary extra labor for a 4.0 in the course was in this group, while the other three students ended up with low but passing course grades.

be measured by whether metacognitive practices, like pausing and tweeting, or reflecting each week on their labors, are being practiced consistently. The nature of that metacognition is what we discuss together as a class or student and teacher, but it need not be factored in my determination of whether my ecology is effective or not.

Furthermore, the nature of metacognition is too varied to have predetermined ideas about what all students should come to. As Tyler Richmond and I discuss, reflection likely is itself a racialized discourse, and may pose contradictions for many multilingual students and students of color who do not share a White racial *habitus* that is typically a part of reflective discourses in college classrooms (142). So I don't want to measure my ecology's effectiveness at getting students to practice metacognitive strategies by assuming a preferred White racial *habitus* in metacognition. Finally, since the grading ecology is a labor-based one, its goals, even ones that promote metacognition, are about measuring how effective the course is at getting students to labor in particular ways, without putting too many constraints on what laboring means. This allows this goal of the ecology to be open-ended and looking for new ways to do metacognition.

GOAL 4: TO SEEK AN AWARENESS OF THE POLITICS OF LANGUAGE AND ITS JUDGMENT.

This goal could easily be considered a part of the previous goal, since its focus on seeking awareness is very close to metacognition, so one might think of seeking awareness of language and its politics as a specific, focused version of metacognitive practices. However, because this goal centers on examining and questioning the power dynamics of language norms, expectations, and judgments that circulate in assessment ecologies, or systems that contain social arrangements of uneven power, I want it to be its own ecological goal. Some of these power relations come from the dominant Discourses used as standards in the assessment ecologies that students circulate in (e.g., White racial *habitus*). Some are from gender, racial, and other social dimensions of those in the ecology itself that intersect with those dominant Discourses.

Regardless of what power dynamics in language and its judgment one wishes to focus on in a classroom, this learning goal requires specific information about such politics, which for some people does not seem the purview of a writing classroom. I don't see how any writing classroom can escape politics, either in terms of pedagogy or curriculum. Politics have always defined language instruction, its history, pedagogies, and assessment practices. I've spent some time drawing out this argument historically and theoretically in classroom writing assessments (Inoue, *Antiracist*) from a racialized standpoint. But I'm not the

first to explain the writing classroom or its assessment in terms of politics. To see the evidence of the politics in writing instruction, we need only look to the discussions on remediation (Rose; Otte and Mlynarczyk; Soliday) that note the historically recurring “literacy crises” that respond to various political exigencies of the time, mostly the presence of people in schools who were not there before, or who weren’t noticed before. We might consider also the discussions around the politics of writing in English departments and colleges (Miller; Ohmann; Crowley; Kynard), or the historically situated politics in testing and assessment (Hanson; Faigley; Elliot; Inoue and Poe; Poe et al.). Or we might look to the discussions in assessment that explain the politics of judgment, who judges are, and where those judges get their values (Guba and Lincoln; Broad; Huot; White et al., *Assessment of Writing*; White et al., *Like a Whale*). No matter the literature one draws on, the conclusion is the same: Teaching literacy means helping students learn about the politics of literacy and how one’s language is judged in the world from those politics.

How does a teacher measure a student’s awareness of the politics of language and its judgment? Some theory about language, race, power, and Whiteness are needed. I will offer a condensed version of some of the discussions I use, mainly Whiteness, but there are others, and it will depend on the teacher’s politics and philosophy of language, as well as the program’s and its curricula. But first, I must say a few words about “awareness,” an important term that is related to reflection and metacognition, and is central to the nature and methods by which this goal’s effectiveness can be understood and measured.

Awareness comes from being mindful and reflective of one’s practices, so this goal is interlaced with the first three above. As Yancey points out in her discussion, reflection is the final step in metacognitive processes. This means that reflective documents are often thought of as a way of articulating what has been learned through metacognitive actions and later reflecting labors. When teachers wish to assess whether a reflective document is adequate or productive, they do not sit on the shoulder of the student, watching them “reflect” as they write. We get the final product, which we then judge as the student reflecting. But it is not the actual reflecting (verb), per se. We are not looking at the act or the labor of reflecting. It is, as Yancey points out, “reflection-in-presentation,” or “a public text representing the self” to a reader (*Reflection* 70). In short, we are reading and judging the noun, reflection, as an indicator of the verb, reflecting, that created the noun in front of us. We don’t really know the labor of reflecting as much as we know the product in front of us as a representation of some kind of previous labor.

Furthermore, most teachers make decisions about their ecology’s effectiveness at getting students to become aware of language and its politics by compar-

ing a present student's reflection-in-presentation on such things against a set of expectations, an exemplar, or a prototype that the teacher holds in their head, three kinds of judging that I explain in Chapter 2 through Haswell's work on holistic judgments. As Catherine Fox argues about critical thinking in critical pedagogies, our typical ways of understanding this kind of goal's effectiveness in writing courses seems to require that we have an idea of what "awareness of the politics of language" is, then judge students against that idea. They have to mimic our (the teacher's) kind of awareness, says Fox (202-03), and this is a Whiteness stance to take as a teacher (201-02).

While I do have my ideas about what an awareness of the politics of language and its judging looks and sounds like, I try not to use that as criteria, or an exemplar or prototype against which to judge students' performances as adequate or not. Instead, I draw on contemplative traditions, traditions that tangibly honor labor as practice done nonjudgmentally—that is, I do not use some kind of "awareness criteria" or exemplar to judge a student's practices of reflection along this goal in order to know how effective my assessment ecology has been. If students are producing reflective documents, recording their labor sessions in their labor logs, then they are doing some form of seeking awareness. How deep or to what degree are questions that have little meaning to understanding my assessment ecology's effectiveness since most of what I can read in a written product is a construction of many other elements in the ecology, in my reading practices, and in the life of that student.

Students can learn and do exactly what they can at this moment. Why would I think that my ecology is less effective because a student who hasn't yet been able to articulate the way Whiteness functions in the judgment practices of his peers, say in their feedback to him, when perhaps she has just recognized Whiteness in her world this quarter for the first time? Is it not enough to see that she is seeking? I know this sounds like I'm saying that anything written down is evidence of this goal if the student says so, but that's not quite it, and there are ways I think one can triangulate. But awareness is a stance in labor, so finding a way to measure the doing of it in labor is the most direct, yet most difficult thing.

To help explain further why I see evidence of this goal in this seemingly permissive way, it helps to understand another assumption about labor in many contemplative traditions. Unlike typical assumptions we make to judge an instance of reflection-in-presentation, contemplative traditions (Barbezat and Bush; Chodron; Hahn; Zajonc) assume that mindfulness, or a reflective stance in a practice, requires the person to first practice, and through practicing, a reflective, mindful stance is slowly acquired, yet never perfected (thus we are always practicing). This stance tends to be called "awareness." While awareness is always a goal, it is not an outcome that tells a practitioner or student that

they've done the practice correctly, that they've arrived and are *aware*. Yet ironically, simply doing the practice is doing it correctly—one has already arrived, one is already here. There is no one way to do it, no set criteria, no exemplar, no prototypes, by which to judge one's practice as aware or not. One cannot do a practice wrong, or perfectly right, for that matter. But one can continue to work at a practice, continue exploring ways to do it, continue seeking to become aware—remember, we are all always becoming. Thus value distinctions, ones that hierarchize instances of reflections, for instance, have no meaning in this “non-judgmental” paradigm of judging the evidence of effectiveness for this goal. Since one can only do a practice like seeking awareness, and explore knowing *that* they are doing it as they do it, it is the *exploring* part that is key. When students are doing the labor, if it is designed right, they are already here, seeking, becoming aware. This is what I look to find evidence of in my ecologies when I measure for this goal's effectiveness.

To measure this goal's effectiveness, I need to gather evidence of the ways students practice seeking awareness of the politics of language, likely imitating at first these reflective stances by posing questions and attempting answers, in order to notice that they are practicing posing questions. Part of this stance of awareness is focusing on noticing while practicing the stance, at first pretending to be mindful and reflective until that pretending becomes more authentic, or habitual—until their practicing becomes *habitus*. But it is unrealistic to expect students to form new *habitus* in a ten-week quarter or a fifteen-week semester, so I look only for the imitating, the initial seeking of awareness in their reflective documents, not to particular conclusions. In several other places, I discuss ways an assessment ecology might meet this goal, as recurring assessment practices that are problematizing (*Antiracist* 237-67), and as Friirian problem-posing, letter-writing activities (“Classroom Writing Assessment”). These reflective documents, as instances of reflection-in-presentation, do offer ways to measure the effectiveness of the ecology at achieving this goal. I look for the presence of the seeking of awareness, not particular practices, conclusions, or insights.

So, the nature of any student's own awareness about the politics of language is that student's business. It simply is not my place to judge whether a student is right or wrong about such things, even if I likely have more information about what they may be trying to learn or understand. I'm still seeking myself. So understanding the ecology's effectiveness along this goal needs to reflect this pedagogical assumption in my ecology. Furthermore, this more humble stance toward my students' seeking awareness of the politics of language is a way to avoid Whitley habits of being the judge, peacemaker, and assuming I know what is right for all students, habits that many have already identified as Whitley habits (Pratt; Fox 201; Frankenberg; Frye 153-54), habits that Fox argues we

must “disarticulate” from our assumptions about criticality in our pedagogies (204). This doesn’t mean that I cannot offer my own ideas about things during the term, but in this chapter, I’m talking about assessing the effectiveness of the ecology’s ability to meet this goal. So judging the effectiveness of this goal in reflections is mostly about noticing when the stance of seeking awareness occurs, without judging the nature of what the student comes to understand or articulate, even though I might make an observation or two about these things to the student during the term.

What this kind of nonjudgmental judgment by a teacher does is demonstrate one counter disposition to those of a White racial *habitus* practiced by most writing teachers. And this kind of judging, not so ironically, illustrates how to apply the theoretical insights that this goal demands students come in contact with. That is, seeking awareness of the politics of language and is judgment is a stance teachers can use to assess the effectiveness of this same goal in our ecologies. Is the seeking of awareness present? How many students are doing this regularly? So, I’m arguing that measuring the nature of the awareness is a different question from measuring its effectiveness in an ecology. Presence, for the most part, means effective, even if it is a rough measure. Measuring the nature of students’ performances that might demonstrate this goal is a qualitative process that might look a lot like Broad’s dynamic criteria mapping, while measuring the effectiveness of the ecology to produce such stances in labor is a matter of finding out if such seeking of awareness is present, regardless of what it looks or sounds like.

Again, I know that many will read this kind of evidence gathering as permissive and loose, perhaps even a way to lower so-called standards. In reality, it is that you’ve accepted one set of assumptions about what it means to measure effectiveness along such a goal, and I’ve accepted another set, a set of assumptions I’ve been attempting to make explicit. I argue that the first set of assumptions, like our schools and academic disciplines, like the dominant English promoted in society, are steeped in Whiteness and promote White language supremacy, while the latter assumptions that center around nonjudgmental judgment attempts to disarticulate those Whitely assumptions in order to form more perfect judgment practices, ones that embody the kind of critical impulses toward valuing diverse perspectives, logics, and Englishes that we say we value.

Let me put it one more way, perhaps to make you uncomfortable if you still are not convinced: If you think that you know best what your students need to know about communication and writing, about what they should be aware of in language and its politics, if you think that you can dictate the standard for such practices, that you are benevolent enough and truly have all your students’ best interests at heart when deciding on and administering your standards of aware-

ness and criticality, then you are enacting a White racial *habitus* that has been one major way schools and society have perpetuated White language supremacy. It is a reenactment of Whiteness, the White savior mentality, and an unnecessary taking on of the White man's burden.

In "White Woman Feminist," Marilyn Frye, drawing on Minnie Bruce Pratt (1984), offers a list of dispositions that equate to Whiteness, or Whiteness dispositions. These dispositions might be thought of as habits of Whiteness, a White *habitus*, and they amount to the following:

- *Being a judge and peacemaker*: a disposition toward giving responsibility and punishments, being the preacher and martyr, taking responsibility and the glory.
- *Self-understood benevolence*: a disposition toward seeing oneself (and other Whiteness people) as benevolent, good-willed, fair, honest, and ethical.
- *Being procedurally ethical*: a strong sense of right and wrong, usually rooted in dispositions toward forms, procedures, due process, and rules as the basis of the ethical; to be good, one acts according to the rules, which is understood as principled.
- *Authority*: a disposition toward running the show, or aspiring toward it, and a belief in one's infallible authority in most matters. (Frye 153-54; Fox 202; my emphasis)

While as the teacher, I try hard not to enact these dispositions of a White racial *habitus* in my responses to my students' writing, I also know that students may still read me as acting in such ways. This goal in the assessment ecology is designed to help the ecology discourage and disable the power of such *habitus*. But the effects are always uneven. I am still the teacher of record. I must "run the show" and distribute grades. Sometimes, I must be the peacemaker and judge, even if those roles do not mean being a preacher and martyr (I may still be read in these ways). Nevertheless, I do not have to act as if my authority is based on infallibility. I do not have to be the "preacher" every time a question is posed to me or the class. I do not have to see myself as a benevolent teacher, etc. I am simply the teacher, a man who is flawed as much as his students are, yet a man who is in a position of power, therefore must act responsibly.

I say all this to say that the effectiveness of this goal in the assessment ecology is just as much about how teachers conduct themselves as how students do. I'm tempted to say that one kind of evidence of this goal's effectiveness in an assessment ecology is how much seeking of awareness of the politics of language and its judgment the teacher is doing? Evidence of a teacher's seeking of awareness might be in things like the teacher's comments and feedback, the absence of

grades and ranking of student writing and work, and a teacher's labor and mindfulness journal that focuses on questions about how that teacher is practicing such a stance of seeking awareness. Are you seeking awareness of the politics of language and its judgment, or have you figured it out already? How long did it take you to get to whatever conclusions you might make about the nature of your current awareness?

Drawing on other theories of Whiteness as well, including most notably Sarah Amed, my students and I have created a set of habits of White discourse that are pervasive in school standards for writing and in readers who judge writing by those standards. There are six habits that are most characteristic of a White racial *habitus*, which leads many teachers to feeling that their own linguistic dispositions, which are not simply linguistic but embodied, as James Paul Gee among others have argued (Young; Inoue, *Antiracist*), are the best standard by which to judge their students. These habits of Whiteness can be used to help students and teacher form practices of seeking awareness of the politics of language and its judgment. My students and I articulate them as follows:

- **Unseen, naturalized, orientation to the world**—an **orientation (or starting point) of one's body in time and space that makes certain things reachable**; assumes (or takes as universal) proximities (capabilities to act and do things) that are inherited through one's shared space; an oxymoronic haunting, leaving things unsaid/unstated for the audience to fill in and contains multiple contradictions (is ambiguous) in how it can be understood; a style of embodiment that is invisible to the person or voice; a way of inhabiting spaces that is comfortable (allows the person to "sink into the space" around the body); the space becomes an extension of the White body and its discourse in such a way that it is hard to distinguish where the White body ends and the world begins; any utterance may participate in this orientation to the world by how the utterance operates in the space (does it sink in?) and what its effects are.
- **Hyperindividualism—self-determination and autonomy** is most important or most valued; self-reliance, self-sufficiency, and self-control are important; individual rights and privacy are often most important and construct the common good; the truth is always good to hear, no matter how painful, good, or bad it may be (each individual has the right to know the truth).
- **Stance of neutrality, objectivity, and apoliticality**—assumes or invokes a voice (and body) or its own discourse as neutral and apolitical, non-racial, which might use some of the other habits to reinforce this neutral and objective stance.

- **Individualized, rational, controlled self**—person is conceived as an **individual who is rational, self-conscious, self-controlled, and determined**; conscience guides the individual and sight is the primary way to identify the truth or understanding; social and cultural factors are external constraints to the individual; meaningful issues and questions always lie within the self; individuals have problems and solutions are individually-based; both success and failure are individual in nature; failure is always attached to the individual and often seen as weakness; control of self is important, as is work and staying busy, or being industrious and productive; unsure how to cope with the uncontrollable in selves, society, or nature.
- **Rule-governed, contractual relationships**—a focus on the **individual in a contractual relationship** with other individuals; focuses on “informed consent” and negotiation of individual needs; individual rights are more important and non-political, whereas socially-oriented values and questions are less important and often political (bad) by their nature; attachment to laws, rules, fairness as sameness; the contractual regulates relationships; little emphasis on connectedness, relatedness, feeling, interconnection with others; individuals keep difficulties and problems to themselves.
- **Clarity, order, and control**—a focus on **reason, order, and control; thinking (versus feeling)**, insight, the rational, order, objective (versus subjective), rigor, clarity, and consistency are all valued highly; thinking/rationality and knowledge are non-political, unraced, and can be objective; anti-sensuality is valued, while there is a limited value of sensual experiences, considerations of the body, sensations, and feelings; a belief in scientific method, discovery, and knowledge; deductive logics are preferred; usefulness and pragmatism are important measures of value and success. (Inoue, “Classroom Writing Assessment” Appendix A)⁵⁸

I introduce and use these habits to help students pose questions about judgments on their writing, and come up with larger questions about the nature of judgment itself. We do this periodically, often after each draft’s assessment letters (feedback from peers) are circulated. If we have time, I push them to ask deeper questions in light of their own power dynamics in their writing groups, or translate their initial insights to other contexts, say another course that they must write in, one that may grade their writing.

58 In the past, my students and I have drawn on the following to help us build these habits of White discourse: Ahmed; Barnett; Brookhiser; Elbow (“Ranking”); Fanon; Fox (“The Evolution”); Frye; Kennedy, Middleton, and Ratcliffe; Myser; McGill; and Ratcliffe.

In the example FYW course I've been using in this chapter, we used these habits of a White discourse to explore differences in peers' judgments made on drafts at two points in the quarter. I asked students to write a letter to me about differences in a few of their colleagues' judgments made on their last draft, looking for ways that the above habits of Whiteness might be informing the way they or their peers read things. To invent the letters, I ask them to build a table with three columns based on their readings of their colleagues' assessment letters. One column has peers' judgments on their drafts (literally copying and pasting of the words from reader's assessment letters to the writer). The next column has the rubric dimension that the judgment just pasted into the previous column is speaking to or about (another copy and paste). And the third column gives the habit of Whiteness that the writer thinks most applies to the judgment. Using these tables, the letters are then written to start a short dialogue about judgment with me.

While I've been arguing that the nature of the seeking of awareness is not as important to measure in order to understand if an ecology is effective at meeting this goal, I understand the need to know the nature of such reflections. T3 offers a typical example of seeking an awareness of the politics of language and its judgment through a look at the way her group members understood one of our rubric dimensions that focused on how we use evidence in our drafts. What made her draft problematic (in a good way) to start with was that it was a poem. The assignment was to dramatically change their drafts from the last version. She chose to change her research essay into two poems, each offering a side of the debate about bullying in school. T3 writes:

One of the main artifacts that all colleagues chose to focus on was, "How does the writer manage or treat evidence of all kinds in their argument or discussion?" When J discussed this artifact in her letter of assessment, she pointed out how she felt that my attempts were unsuccessful. "I felt this dimension is where you lacked in your poem. In the directions it said that we need at least 3 sources of evidence in your poem, I struggled even finding one. Early in this draft I saw that you tried to tie the movie "Mean Girls" into it by saying "All I could think about is how conflicts were handled in movies." . . . I see what you were trying to do here by using the movie as a source, but it did not work for me" (J). I struggled reading this line because I thought that my letter beforehand would clear any confusion as to why there were not any sources but at least I had tried. I thought that if I had pointed out my

own errors before hand, then maybe the readers would be a bit more understanding that incorporating three blatant sources into what I was writing was extremely difficult. Then Jamerika mentioned, “I do not think it necessarily even needs sources, but just for our requirements” (J). After reading this I thought, “Wait.” J kind of contradicted herself in a way [. . .] Which one does she want me hook onto? Making my sources better? Or realize that they are not even needed? S, however, contradicted with what J had first mentioned about my sources. “I understand that as your draft was a personal invention, there’s really only minimal need for sources. Besides the necessary ones you included of course, sources really do not make much sense in this specific context” (S). This makes sense to me. I felt that when I first started writing that sources were not really necessary, even though it is a requirement. But if I want a good grade, then I need to follow those specific labor instructions, right? [. . .]

So far here is how I have interpreted the contradictions: 1: “Your attempt at including evidence was not successful. You should try X, Y, and Z to make it more efficient—BUT, your poem does not need sources anyways.” 2: “Since this is a poem, sources are not even needed.” [. . .] I am at a crossroad.

[. . .] How can you successfully include sources in something that is not a traditional essay?

What I find exemplifies this ecological goal the most in this letter is the way in which T3 maintains a questioning attitude toward her colleagues’ judgments, trying to hear the way in which each is reasonable. She comes to a good question that really is rooted in a sense of the politics of language, even though she cannot articulate those politics here. How do you include sources in poems (in a way that all readers will recognize and accept)? T3 is asking good questions that come from judgments on her draft. While she cannot yet articulate a question about judgment itself, she is asking about expectations and norms, which is fundamentally a question about the politics of language and its judgment, audiences and expectations, genres and conventions. As the reader of this letter, I can respond by helping her articulate these nascent questions as political ones, ones about who has the power to determine the answers, under what conditions, and to what effects in society or schools.

In T3’s table, she links J’s contradictory judgment, but one that ultimately

sides with the draft's requirements (that assumed a conventional paper), to the White habit of an "Individualized, Rational, Controlled Self," explaining that "I mentioned in my letter that I struggled with including sources throughout the piece, and my subtle attempt of including a source was not viable for J." T3 associates S's judgment that centers on sticking to the genre expectations of poems with the habit of "Hyperindividualism," which she explains in this way: "[s]ince I did a personal poem, sources aren't necessarily needed. This contradicts J's point." Because the table doesn't ask her to come to some conclusion, just record observations, it helps T3 seek awareness of the politics of language, even if she doesn't articulate clearly these politics in the letter—that is, articulate the contradictions in judgments that she raises as uneven power dynamics and differences in expectations or norms that are enforced in most school assessment ecologies through the use of grades and a teacher's authoritative feedback.

The next step would be for T3 and I to talk about how these judgments, while they participate in such habits of Whiteness like an individualized, rational, and controlled self, and hyperindividualism, are not engaging in White language supremacy. Why? It is simply that J is a Black female colleague (S is a White male)? White language supremacy can be enacted through a body of color. Both do not engage in White language supremacist judgment practices because our assessment ecology keeps that from happening. How such judgments are circulated and their effects on those in the ecology matters to seeing such White habits as enactments of White language supremacy. The draft is not graded or ranked, so even if judges have engage in such White habits in feedback, as J and S seem to have in differing degrees, they are not used against T3 to force her to do something she isn't interested in doing in her writing or punish her for not doing it in the first place. She has a real revision choice, and the labor-based ecology helps her seek awareness in a safe way. This context makes the mere presence of such reflection in the letter an indication of the ecologies effectiveness along this fourth goal.

The measure I use to assess the ecology's effectiveness at this goal, then, is not what T3 and her colleagues come to understand or articulate about White language privilege. Doing so would be a Whately act itself on my part, since I'd have to use my own notions of what students should come to understand about the politics of language, regardless of their own subject positions or goals for our course. Instead, I look for the seeking of understanding through the questioning of judgments, which she and most of her colleagues do.

In retrospect, I can see one way that I might be able to further set up the assessment ecology so that all students demonstrate such a seeking of awareness. I might add one column to the brainstorming tables that all students must complete before they can write their problem-posing letters to me. This new column

might ask them to pose a question or two that each row of information reveals to them about the politics of language, or about the judgment made in that instance. As an example, consider a portion of T3's table that she submitted with her letter to me, which I've revised to include a new column (far right, "Question(s)" column) that poses a question or two about the politics of language and its judgment that might come from considering the information in that row.

Table 7.5. An example of a table that shows evidence of seeking an awareness of the politics of language and its judgment

Colleagues Judgments	Rubric Dimension	Whiteness Trait	Question(s)
<p>"In the directions it said that we need at least 3 sources of evidence, and in this poem, I struggled even finding one . . . I saw that you tried to tie in the movie "Mean Girls" . . . I see what you were trying to do here by using the movie as a source, but it did not work for me" (J).</p>	<p><i>How does the writer manage or treat evidence of all kinds in their argument or discussion?</i></p>	<p>Individualized, Rational, Controlled Self—I mentioned in my letter that I struggled with including sources throughout the piece, and my subtle attempt of including a source was not viable for J.</p>	<p>What is "rational" or "controlled" about explicitly using a source in a poem, or even an essay? Where do these norms or expectations that J identifies in our assignment come from?</p>
<p>"I consider this structure to create clarity, effectiveness, etc. I understand that as our draft was personal invention, there's only minimal need for sources. Besides the necessary ones you included of course, sources really don't make much sense in this specific context" (S).</p>	<p><i>How does the structure of the draft create clarity, effectiveness, or persuasion for the reader? // How does the writer manage or treat evidence of all kinds in their argument or discussion?</i></p>	<p>Hyperindividualism—Since I did a personal poem, sources aren't necessarily needed . . . This contradicts J's point.</p>	<p>Where does S get his expectations about poems and sources? Why value those expectations here, ones that are "hyperindividualistic," that value mostly my voice in the poem? Does our class allow us to value them? Would another class not value them?</p>

The exact questions are not that important for determining the ecology's effectiveness at producing this goal. Instead, it is the presence of the questions and how they link to the data in the other cells in the row of the table. It's a sim-

pler way to see effectiveness without overly determining in a Whately way what students should come to. Of course, this is, as I've discussed throughout this chapter, only the way a teacher would determine how effective their labor-based grading contract ecology has performed along this goal. I am not speaking to how a teacher might respond to such student reflections, or what discussions classes might have from them, or what we might learn afterwards about our ecology by looking at the nature of students' reflections on the politics of language.

Table 7.6 offers some kinds of evidence that can be useful in determining a labor-based grading contract ecology's effectiveness at encouraging this goal.

Table 7.6. Evidence of awareness of the politics of language and its judgment.

Evidence of Seeking Awareness of the Politics of Language & Its judgment	Questions to Ask of the Data
Problem-Posing reflective activities (reflection-in-presentation)	How many students complete the problem-posing activities? Do most students pose questions about judgments of their writing that look to understand how each judgment is reasonable? How many questions do students pose on average?
Tables or other observation-based heuristics	How many students complete the heuristic activities? Do most or all students make observations about White habits of discourse in judgments of their writing? Can students make observations about various <i>habitus</i> that inform judgments on writing? How many questions on average do students make?
Problem-posing journals (teacher or students)	How many journal entries does each student complete? How long are the entries on average? Do they get longer as the course goes on? Can students form observations and questions about the politics of language and its judgment in most journal entries? Do students engage in questions about White language privilege or other <i>habitus</i> in language and judgment in most or all entries of the journal?

The effectiveness of an assessment ecology along this goal can be argued often by simply counting how many students, and how often, engage in questioning the nature of the judgments of their or others' writing. As can be seen in the right

column, much of how I determine such counting is by determining: (1) how many students are able to complete most or all of that kind of labor; (2) how many questions students can pose (the more present, the higher degree of effectiveness of the ecology); and (3) how many students tried to employ the theories of Whiteness and White language privilege to their labors?

When measuring for this goal in a course, a teacher should be mindful of who their students are and where they come from. Many privileged, White students will have a hard time accepting something like White habits of discourse that contribute to White language supremacy.⁵⁹ They will hear it as “reverse racism,” or a personal attack. This is not a reason not to engage in such questions. In fact, it is the reason to do so. In most classrooms in which I’ve taught, the students are mostly students of color and working-class. Many are multilingual. So this goal is an easy one to discuss with such students, many of whom already feel oppressed by educational systems that have punished them for their language use. Such oppression usually gives a person some automatic awareness of the politics of language.

GOAL 5: TO MAINTAIN SOCIALLY JUST CONDITIONS FOR LEARNING BY ENSURING EQUITABLE OPPORTUNITIES TO RECEIVE ALL FINAL COURSE GRADES POSSIBLE.

As I discussed in Chapter 2, Rawls’ notion of justice as fairness dictates that everyone in a society should have a “fair equality of opportunity” (43). In assessment ecologies in writing classrooms, this means that in order for an assessment practice to be socially just, all students in the ecology should have equal access to all of the opportunities that afford success in and from the ecology, which includes access to all of the possible course grades. Con(De)fining precisely what students should learn in a literacy learning course by judging students against a single standard or outcome, then, will exclude some students from having full access to all possible opportunities for learning, success, and grades. This does not make for a socially just assessment ecology.

As I argue at the opening of this chapter, when used to measure success of students or the effectiveness of a course, outcomes are a socially unjust means to do so, because by their nature, they are exclusionary. Measuring for this final course goal helps me attend to ensuring equal opportunities for grades to all students in my courses. While grades are not learning, they do directly affect students’ latter opportunities in and out of school, so as an indication of future

59 One way to prepare some White students for discussions of the racial politics of language is to begin with readings on White fragility by Robin DiAngelo (“White Fragility”; *White Fragility*) and cultivating “brave spaces” for such discussions by Arao and Clemens.

consequences and available opportunities in and after school, they are important. Furthermore, grades affect the psychology of students, their attitudes and confidence levels in present and future tasks, which is another way of saying they affect students' abilities to do the labors of learning in school. Relatedly, Inman and Powell have argued that grades in writing classrooms have important affective dimensions associated with them that may determine students' abilities to learn or succeed in a classroom. They often are symbols of desire, "a reassurance of rigor" (42), or a symbol of regulation that controls behavior (44). So because they exist already in educational systems, grades are still important, perhaps through a haunting White presence outside contract graded classrooms.

Beyond my application of Rawls' theory, it is important to understand Iris Marion Young's "social connection model" for responsibility that structures social justice in a society, and in my labor-based contract grading ecologies. Young's theory is predicated on mid-twentieth century observations about individuals' social responsibilities to all of the other members of a society. She explains that "members of a whole society collectively bear responsibility for taking care of one another's old age, health care, and children, and for keeping us out of poverty" because we are all interconnected. She says, "people owe one another a certain measure of reciprocal care because of these interdependencies" (9).

In order for societies to meet these reciprocal responsibilities, we have to have social and other structures in place in order to prevent inequalities and injustices from occurring (33-34). Young's explanation of her social connection model for a responsible society works just as well for classroom writing assessment ecologies. In fact, one can hear the writing classroom in Young's description of the model:

The social connection model of responsibility says that individuals bear responsibility for structural injustice because they contribute by their actions to the processes that produce unjust outcomes. Our responsibility derives from belonging together with others in a system of interdependent processes of cooperation and competition through which we seek benefits and aim to realize projects. (105)

So like antiracist writing assessment ecologies (Inoue, *Antiracist* 93-104), Young's social connection model for responsibility assumes that we are all interconnected because we are elements in larger structures in society that are interdependent, which makes us structurally interdependent. This interconnectedness as an explicit aspect of a classroom grading ecology is vital to antiracist projects (Inoue, *Antiracist* 104), and to assessment ecologies that attempt to help students form critical stances of judgment. We need each other if we all are to

have fair equality of opportunity, and to understand that our perspectives and views are located in a universe of other perspectives and views. In order for any one person to learn and succeed, that individual requires a web of others to help them.

To help me design and assess how effective my ecology is at achieving this goal, I ask myself a simple question about my syllabus and the grading ecology I've designed, which I later ask my students at week 1 and 6 (in a ten-week quarter system) when we negotiate the grading contract together: *Does everyone in the class have an equal opportunity to get all of the possible course grades available?* Is this assessment ecology fair enough for everyone? If we think it isn't, then what rules and structures in the contract do we need to adjust or change? What structures and assignments need to change in the course? Checking students' felt sense of the fairness of a labor-based grading contract is one way to measure and ensure this goal. If students feel that the contract is fair, and they've had at least two opportunities in the quarter or semester to negotiate the terms of the contract, then there is a high probability that most or all will find it fair enough. And such negotiations are important in measuring the effectiveness of this goal. That is, another measure of this goal is whether the ecology offers these two moments in the quarter or semester to (re)negotiate the contract's details.

As Chapter 4 explains, these two moments of negotiation are important. Getting students' assent or agreement about the fairness of the contract in the first week may not be enough to ensure a contract's acceptable level of fairness. In fact, one should double-check this, since students may feel coerced into agreeing to the contract because most of the time FYW courses in college are required. Many students may not feel they can just drop the class, or contest ideas in the contract. This also means that beyond having a second, midpoint, renegotiation, I look also at grade distributions of the course after it's over. While some students may not choose to pursue higher grades in a labor-based system for a lot of reasons, there still should be more Bs and As than in conventional bell curves (or a "normal distribution"), if one assumes that more students than not will attempt more labor to get higher course grades.

To understand how I use grade distributions to help me determine how effective my grading ecologies are at this goal, it's important to understand a bit of statistics. I am by far not an expert at statistics, but one doesn't need to be an expert in order to understand grade distributions and make observations about equal opportunity in the ecology. A statistical normal distribution, or bell curve, is a distribution (of grades in this case) in which about 68% of all the grades are within one standard deviation from the mean (the average grade). This means that about 68% of the scores plotted on a graph fall on either side of the median score (the exact middle score) on the graph. About 95% of all grades fall within

two standard deviations from the mean, and 99% of all grades are within three standard deviations from the mean. Figure 7.2 illustrates a typical bell curve and where the percentages of the data are relative to the median score (designated as μ in Figure 7.2).

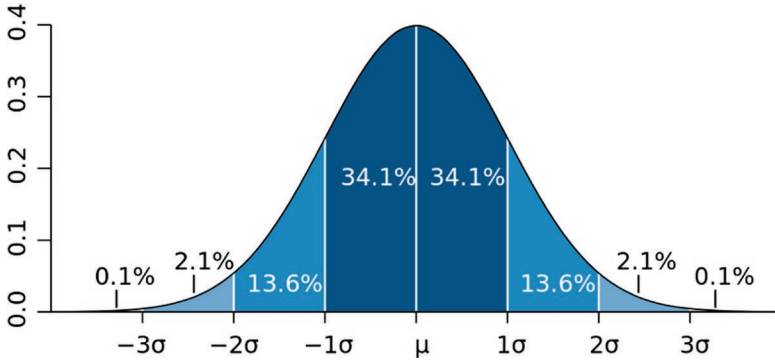


Figure 7.2. A standard bell curve, or normal distribution.

While there is much evidence to argue that in human populations many measurements do naturally result in such bell curved distributions, for example the height of men and women. However, the instrument of measuring is vital in having both highly valid and reliable results by which to create such distributions and call them accurate assessments of diverse groups of people. One can establish reliable measuring protocols and instruments to determine the height of a diverse range of people. No one disputes the length of an inch or a centimeter. The grades in writing courses, or grades on writing assignments, however, are a more difficult and disputable measurement, and the instruments (scoring guides, standards, rubrics, etc.) used to determine such grades are highly diverse and idiosyncratic since they come from and are applied by the human brain. It is difficult to establish anything like reliability in any set of grades on writing, which past research has shown.⁶⁰ Furthermore, accepting scores or grades that then get used to produce a grade distribution means we accept whatever standards of writing were used to produce those grades.

I say all this to qualify what I'm about to offer. I do not mean to suggest that course grades tell us the fairness of a grading ecology in a writing course, but they do offer some indication of unfairness. One assumption that Young makes about all social systems, such as conventional grading ecologies in writing classrooms, is that they are inherently unfair to some students, thus we must work at

⁶⁰ As mentioned in Chapter 6 under the question of "don't students want or need grades . . ." there is lots of research that illustrates the unreliability of grades on writing in high school and college settings.

not being unfair—at not simply accepting the status quo, the standard operating procedures of grading in schools.

When I use grade distributions to help me determine how effective my ecology has been at achieving this goal, I compare my data to a simple rule. The rule is that if my course's grade distribution can show that most students who abide by the labor-based grading contract get good grades (B-/B range), then my ecology has ensured a reasonable degree of equal opportunities. While in most classrooms, a single writing teacher can determine this without the need of calculating mean grades and standard deviations (SD) of those distributions, doing so can help programs with many courses, where such eyeballing of grades and students is much harder to do. But calculating the mean grade and SD for a course can also provide a numerical way for a teacher to track their courses' assessment ecology's effectiveness at achieving this goal over time. For these reasons, I use median grades and SDs. In my context, I feel it is a reasonable goal for me to expect that 75% of all my students should achieve a course grade of 3.1 (B). This means at least fifteen of the twenty students who normally fill my FYW courses should be able to meet the contract or exceed it. That's my numerical target for an acceptable effectiveness level at this ecological goal.

I am a data hoarder on my courses and students. I have always wanted to have evidence to understand what is happening in my classrooms, so I have my course grades from two decades of teaching first year writing, and those distributions tell a story in three parts that is instructive here. Table 7.7 shows the mean grade and standard deviations for my course grades in three different periods of my teaching. These periods of teaching correspond to the way I structured my FYW courses' grading ecologies. The first two rows of data are the first two periods, and rows three and four make up the last period. The rows in the table are: (1) FYW courses during the academic years of 1998–2000, where I used conventional grades and standards; (2) FYW courses between the academic years of 2001–2004, where I was experimenting with grading systems, but not using labor-based grading contracts; (3) FYW courses between the academic years of 2005–2012, where I used labor-based grading contracts; and (4) FYW courses between the academic years of 2014–2016, where I used labor-based grading contracts at UW Tacoma.

I separate rows three and four because the nature of the courses are different enough that they are worth distinguishing. Row three represents courses in a sixteen-week semester system, while row four represents courses in 10-week quarter-based school system, my current university in the state of Washington, whose campus is mostly working-class students of color. The length of time in the ecology does seem to matter more than other ecological elements. As Chapter 3 discusses, time matters when labor is the primary measure of value and worth in

an ecology. I also acknowledge that Row one represents courses from a two-year college in Salem, Oregon with a mostly White, working-class student population. Row two represents courses from a primarily White, four-year university in Washington State. Row three represents courses from a four-year regional university in southern Illinois that served a large working class, African-American population, as well as a four-year university in central California that was an Historically Hispanic Serving Institution, whose students were mostly working-class.

Table 7.7. Mean grades and standard deviation of final course grade distributions for My first-year writing courses

	Mean Grade	Standard Deviation	# of Ws or Is (percent of total)
Conventional FYW Courses (1998–2000)	1.79 (D+/C-)	1.41	15.26%
Experimental FYW Courses (2001–2004)	2.92 (C+/B-)	.79	5.56%
Labor-Based FYW Courses (2005–2012) ⁶¹	3.18 (B)	.40	1.41%
UWT Labor-Based FYW Courses (2014–2016)	3.32 (B+)	.44	7.14% (1.79%)

The means and standard deviations in Table 7.7 do not calculate students who withdrew, dropped, or received incomplete grades, but it is worth noting the differences in each group. I've included the percentage of the entire group of W and I grades (I do not have records for those students who dropped my courses) given in the far-right column. This is one indicator of the level of unfairness in a grading ecology. When 15% of all my students registered for my FYW courses withdraw or got an incomplete (as in Row one), it suggests a higher degree of unfairness in those conventionally graded ecologies than the other ecologies represented in the table, which used different grading ecologies. I should acknowledge that there are lots of reasons for a working-class, White, mostly first-generation group of students to withdraw or get incompletes in a course. There are limitations to such data, but it is a good place to start inquiring about an ecology's effectiveness at this goal. I don't want to ignore such data, even if I might have ways of explaining the numbers.

61 I should note that from 2005 to about 2009, my grading contracts, as I discuss in Chapter 2, were hybrid contracts, but still mostly labor-based (up to the B course grade).

The differences in Ws and Is move down quite a bit as I move to experimental assessment ecologies, and finally labor-based ones. The higher percentage of Ws and Is in my UWT labor-based ecologies (7.14% in Row four) is likely skewed by my first FYW course taught at UWT in my first quarter returning to a quarter system after about fifteen years away. Both Row one and four represent ten-week quarter system courses. So in parentheses in Row four, I've included the number of Ws and Is from this group without that first class. What also skews this number is the fact that it has less data than the other rows because of the fewer FYW courses I generally teach annually at UWT.

As Table 7.7 shows, the mean grade (based on a 4.0 scale) went up over time in my FYW courses, meaning as I moved to labor-based contract grading ecologies in my FYW courses my mean grade, the average grade given, moved from a D+/C- (1.79) to a C+/B- (2.92), to a B (3.18), then to a B+ (3.32). The total movement of the average grade given was more than a full letter grade, but the differences in standard deviation (SD) I find most interesting. SD shrinks as my grading ecologies move to a labor-based system. SD tells us how much variation there is in the grades, or how dispersed the grades are in the distribution of all grades in that category on the table. So if I want to be most effective at this goal, I would need as evidence a high mean grade and a small SD, or a small spread. This will mean more students hover closer to that mean, whatever it is, if the SD is small—that is, if the spread of grades is not that dispersed.

According to Table 7.7, when I used conventional grading ecologies, about 68% of my grades were 1.41 grade points from the mean grade of D+ (1.79). In effect, 68% of all my grades were between a 3.2 (B-) and a .38 (D-/F). This also means that some within one SD of the mean grade in those classes failed the course. On a 4.0 scale, a 1.41 SD could be considered slightly dispersed, or a wider spread than expected. To give you another way to see this, in Row one's courses, if the grades were evenly dispersed, then in a class of twenty, about seven students received grades between 1.79 and .38. At UW, where we use a numerical 4.0 scale to record final course grades, 1.79 is in the middle of the C- range (1.8-1.5). This means that most students in my courses did not get high grades. In fact, if we consider the lower half of the C- range as failing, then 50% of all my students received non-passing grades. This is not an effective grading ecology at achieving equal opportunities for success.

Compare the first row to the groups with the smallest SD, the FYW courses in which I used labor-based grading contracts, or Rows three and four. In Row three, for instance, 68% of the students received between 3.58 (B+/A-) and a 2.78 (C/ C+). Most students received average to higher grades, and all passed the course. A failing grade is a 1.4, so even students within two SDs of the mean passed the course with at least a 2.38 (C+) final grade. Two SDs accounts for about 95% of all

students. This meets my high standard for effectiveness along this ecological goal. In my current FYW courses at UWT, represented in the fourth row, similar effectiveness numbers can be seen. Sixty-eight percent of students received between a 3.75 (A-) and 2.88 (high end of the B-). And because of the high mean grade, only about 2% of students received final grades lower than 2.0 (C). These are even better numbers, thus show a high degree of effectiveness.

If one measure of unfairness is a relatively large number of W and I grades, then another measure of unfairness might be the SD of the final course grades in a class. As you can see from my discussion, it isn't as simple as saying the smaller the SD, the fairer the ecology, since a small SD in the first row of data in Table 7.7 might still show too many students failing or struggling unnecessarily if the mean grade is low. But having a smaller SD in a course that has an average grade of 3.18 or 3.32 does appear to be a fairer ecology, if we accept a higher degree of passing students as an indirect indicator of equal access to all course grades, and really we mean equal access to higher grades, since anyone can get a low grade by doing nothing.

Based on the average grade for my labor-based grading contract ecologies in those rows (3.25) and their small SD (.42), my labor-based contract grading ecologies can be seen as highly effective at ensuring equal opportunities to receive all possible grades since about 98% of the grades given, which accounts for two SD from the mean grade, were passing grades, or 2.42 (C+/B-) or higher. In other words, the vast majority of students had access to the higher grades on the scale. Add to this the fact that these same labor-based ecologies had fewer numbers of W and I grades, dramatically fewer than my conventionally graded students. This makes for compelling evidence of a high degree of effectiveness of labor-based grading contract ecologies to ensure equal opportunities for high grades.

But looking at one's grade distributions and considering how close they are to normal distributions may not tell a teacher enough to know how effective their assessment ecology has been at ensuring equitable opportunities for all grades to all groups of students. Who gets what grades and under what conditions? Do students of color, for instance, choose not to pursue more labor than White students, generally speaking? Do some groups of students have more time constraints during the semester than others in my grading ecologies? These are difficult questions to answer that can only be figured out once a course begins, and the class meets. I find labor logs and labor journals help in my looking for such problems. But we might also look at mean grades and SDs within important racial formations.

Table 7.8 offers these data by racial formations in my UW Tacoma courses between 2014 and 2016, because I have access to these data. The data is limited,

because I only teach between one to two FYW courses a year. The table does not represent a sample. Table 7.8 represents all my FYW students. Because my classrooms are racially diverse, there appears to be enough data on each racial formation to calculate means and SDs, with the possible exception of the Latinx formation, as suggested by that group's higher SD in the distribution of grades. One measure of a grading ecology that offers all racial groups equal opportunities for all course grades is a high mean grade and a small SD that is consistent with the other groups measured. In the UW system, a 1.5 is the lowest course grade that is passing in any course (C-). No student in my classes has received such a grade. In other words, if you got to the end of the quarter, you got at least a 1.6 course grade. A few have withdrawn from the class because they would not be able to get such a final grade. The number of these students is shown in the far-right column, with a percentage of the group in parentheses.

Table 7.8. The Mean grade and standard deviation in grade distributions of four racial formations in my FYW courses at UW Tacoma

Racial Formation (total students)	Mean Grade	Standard Deviation	Total # of Ws
Asian (12)	3.33	.43	1 (8.3%)
Black (11)	3.12	.51	1 (9.1%)
Latinx (7)	3.30	.87	0 (0.0%)
White (26)	3.40	.31	2 (7.6%)

How consistent is the mean grades and SDs of each group compared to the others? Are they all about the same? And do they all equate to the same number or percentage of students in each racial formation passing the class and achieving the same kinds of grades? If they are similar, then one could argue that my grading ecology is antiracist because it offers equal opportunities for success (and failure) for all racial formations in my classes. As you can see from Table 7.8, this is mostly the case. The slightly higher SD for the Latinx group could be due to its lower number of total students, since one or two students out of seven would skew this number, but I would consider this inconclusive. And yet, the Latinx numbers are not that far off. Black students still do the worst of all formations, with a slightly lower mean grade and wider spread in the SD, but even in this group, 98% (which accounts for two SDs from the mean) received a 2.1 (C) or higher final course grade. The Latinx group is also still within this same threshold, with 98% of students getting a 1.56 (C-) or higher course grade. Generally speaking, these are consistent numbers that all show 98% of students with access to higher grades and succeeding in our labor-based contract grading ecologies.

These data show me a grading ecology that is socially just, offering equitable opportunities to all for the highest grades.

CONCLUDING WITH TWO CRITICISMS

I'd like to conclude this chapter by addressing two criticisms that could be leveled against labor-based contract grading ecologies that have direct bearing on how effective some may find them. One important measure of effectiveness, no matter the ecological goals being measured, should be student reactions to the grading ecology. How do students feel and experience the grading ecology? There are a number of ways to define such a measure. Typical ways might be through end of course anonymous surveys of students that ask any of the following:

- How happy were you with the grading contract?
- How satisfied were you with the grading contract to create a fair grading environment?
- How helpful was the grading contract to your learning in the course?
- How effective did you find the grading contract for your purposes in the course?
- How effective was the grading contract in allowing you to take risks in your writing, or to learn the kinds of writing skills and practices you were hoping to?
- Do you prefer the grading contract for writing courses over conventional grading?

While each of these questions define student reactions differently, they get at much of the two criticisms that I think need consideration.

Both of the criticisms are tied to the labor component of the ecology. The first concerns criticisms of contract grading offered in Spidell and Thelin's study of hybrid grading contracts that looked at student reactions to them, published in 2006. The second is a criticism of writing classrooms that overly regulate the labor and activities of students by promoting what Soliday and Trainor call "audit culture," which is an anthropological term. Soliday and Trainor's study also uses qualitative measures of students' experiences in classrooms. I find both studies to be important works that help explain why labor-based grading contracts, at least in the ways I try to enact them (and discuss them in this book), avoid the problems of student dissatisfaction with any grading ecology, thus it's important to hear that I assume students should feel good, satisfied, and accepting of whatever grading ecology they are in. So these two criticisms are about effectiveness, since part of effective grading ecologies, for me, is how students experience them and react to them.

I should note that in another place I critique the raceless methodology used in the Spidell and Thelin study, and demonstrate a way to have a racial methodology in order that one can see and make arguments about the racialized dimensions of one's data ("Racial Methodologies"). Despite this problem with the study, Spidell and Thelin offer important insights to those contemplating the effectiveness of their own hybrid grading contract ecologies. What Spidell and Thelin's study shows is how mostly working-class White students at a Midwest university found a hybrid grading contract much like Danielewicz and Elbow's (discussed in Chapter 2) in writing courses. They make the following findings about their hybrid contracts:

1. students resented not getting grades, or not having their work quantified by grades (41);
2. the contract tended to create more anxiety over and resistance to the work in the course (43-44);
3. the contract made typically high-performing or highly motivated students feel an unfair "leveling effect" (i.e., it didn't seem fair that their hard work over drafts was counted the same as someone else's, less-than-hard work) (45);
4. students wanted more input into the construction of the contract (48);
5. the contract made the course more difficult than necessary (48);
6. there were no consistent attitudes or reactions to the contract (50-52); and
7. students had a difficult time seeing the difference between hybrid contracts and conventional grading systems (52-54).

Spidell and Thelin do not argue to writing teachers that hybrid contracts are not worth the time, instead they argue for a cautious use, one that is tempered by working closely with students, negotiating the terms of the contract, and helping engage in dialogue about their work along the way. In short, they suggest writing teachers listen to their students. A good principle. The six concerns above, which equate to the study's conclusions and title, "not ready to let go," might be seen as a criticism of the labor-base grading contracts I discuss in this book, but the ecology from which they spring are not the same as those I've been discussing. In fact, labor-based grading contracts address most of these concerns, and one could say they also offer unique conditions for deeply attending to our students, which can also address several of these student concerns.

In other places ("Grading Contracts" 88; "Racial Methodologies" 134), I show the way students from various racial formations in labor-based writing assessment ecologies at California State University, Fresno, show acceptably high levels of happiness, helpfulness, effectiveness, satisfaction, and preference (five of

the kinds of student reaction questions listed earlier in this section) for mostly labor-based grading contracts in writing classes.⁶² Students in labor-based contract grading ecologies do not seem to hold resentment toward the lack of grades (Concern #1 above) when they have ongoing conversations about grades, about what they mean, how they have been used against them, and their effects on students' motivations and feelings about themselves as writers. In my ecologies, our parallel discussions of compassion and caring discussed in Chapter 5 help students lose much of what Elbow describes as a "hunger for ranking" ("Ranking" 190-91, 197), which is at the heart of this first concern. But it isn't the only thing creating such a reaction to the Spidell and Thelin's hybrid contract. It is also the contradiction in philosophies of assessment at work in the hybrid contract, which I've discussed in Chapter 2.

The hybrid nature of the contracts in Spidell and Thelin's study likely had something to do with the student concerns voiced in #1, #2, #3, and #7. The fact that on many writing assignments, usually more formal assignments, the teacher would still determine the quality of the writing, as explained in their contract for an "A" grade (64). This creates a contradiction in how some grades are determined (those above a B). Leaving the teacher's judgments of quality in the grading ecology as a way to determine a crucial distinction, that between B grades and A grades, means that students still have to care about grades on every writing assignment. They might not be worried about failing so much, but if they care about getting a better grade, then they surely care about those quality judgments, the very thing that contracts are meant to mitigate. Letting the teacher's judgments of A-quality work circulate in the ecology creates this contradiction when part of the contract determines some course grades in other ways. These competing logics that produce signifiers of success and achievement (one based on labor for grades up to a "B" and one based on a teacher's judgments of quality for those above the "B") will create more anxiety and resistance to the work of the course (Concern #2)—it is easy to feel like you don't know how well you are doing when grades are left off writing until the very end of the term, when it really matters.

This same dynamic in student responses to hybrid contracts can be seen in Inman and Powell's more recent study, where students still had concerns about "where they stand" in the class (41), which the authors attribute to a deep-seated desire for grades. This desire stems from students senses that grades are "signifiers of how much work remains to be done," which is connected to their identi-

62 Because these conclusions come from a program with many different teachers using the same grading contract template, I cannot say that all instances of the contract were purely labor-based, likely some were hybrid. However, the mandatory training the program gave at the time to teachers on the use of contracts was labor-based.

ties in school (42). Inman and Powell argue that grades have become “affective carriers of emotion” for students, “markers of achievement” that teachers cannot (and perhaps do not want to) ignore. However, like Spidell and Thelin’s study, Inman and Powell neglect to consider the hybrid nature of their contracts, and how the presence of both labor-based and quality-based decisions about progress and achievement in their study’s courses might confuse students and make them anxious about how well they are doing. Neither study considers the real possibility that withholding quality judgments in an ecology until the very end of the term, when the A/B distinctions are made, would make many students feel a need for grades because they know that judgment is coming and is important still.

Thus hybrid contract ecologies create a contradiction and perhaps anxieties around it. In effect, students receive contradictory messages. On the one hand, they are told the contract is here to help you forget about grades and take risks, to just write, and do all you can to develop yourself as a writer. On the other hand, the way higher course grades are determined is the same way they always have been, by a teacher’s judgment. No wonder students had a difficult time distinguishing their hybrid contract from conventional grading (Concern #7) in the Spidell and Thelin study.

According to Inman and Powell, however, it is students’ affective-based desires for grades, which are needed references to their status and identity that causes such problems in students’ reactions. They give a student response as an example of this need for a reference-grade as a marker of status: “[I] did not understand if [I] was average or below . . . I really do not know where I stand in this course” (40). The authors explain these comments by saying that the student understands her standing in terms of a grade, which also tells her if she is average or below (40). They suggest that it is the history of grades that creates such desires and affective dimensions within students (36-7), but they do not mention how the presence of even just A-grades that are determined differently than those below them, can cause such confusions and desires. In fact, Spidell and Thelin gesture to this problem, saying, “[w]e wonder, then, if, adhering to letter grades for different types of writing throughout the contracts—being clearer about expectations—somehow produced the discomfort or uncertainty in these classes” (42). In contrast, Labor-based grading contracts are clearer for students because they do not contain the contradiction that hybrid ones do.

It may be obvious to some that Concerns #4 and #7 are not issues in my labor-based grading contract ecologies. We have a full week and half of initial negotiations of the contract that begin with understanding what the contract is saying and how it is different from conventional grading systems, and renegotiation processes at midpoint (explained in Chapter 4). I also provide different labor instructions for all assignments, as well as ask students to be continually

mindful of their labor practices. These practices keep students from confusing our assessment ecology from conventional ones (Concern #7).

But as Spidell and Thelin remind us, hearing from students directly is vital in understanding their reactions. In my own ecologies Concerns #1, #2, #3, #4, #6, and #7 do not appear to be a problem, even up to the midpoint renegotiation processes, which begin with reflections on how the contract has treated them, how they are doing in the course, and how fair they find the contract now. In the example course I've been using in this chapter, no students in the class said the contract was unfair at the end of week 5. In fact, all found it fair and had positive observations about it. Student B3 (a White female) in the Bottom performing group from Table 5.1, offers a typical reflection on the contract at this midpoint:

At the start of the quarter I didn't really appreciate the late to class/missed days policy because it just sort of seemed unfair if there were emergencies. My dad has leukemia and gets unbelievably sick quite often, and when he is in that sort of condition he needs someone there to take care of him, my step mom works full time and often can't get out of work so I'll stay and make sure he's drinking water and getting rest. I understand now that my education should be a priority and I don't really like missing class anyway, but I have worked out with my family how my step mom can wait for me to get home and then she will go to work and I'll spend the rest of my day taking care of my dad. In my eyes the contract still seems really fair, with class only being three days a week it's pretty easy to get over being sick quickly or working it out with our professor.

The contract still wonderfully reflects the course goals, over-time I have managed to make friends with the people in our class and I enjoy hearing their constructive criticism of my work, I genuinely value their opinions. While in high school I never really talked to many people just because I was shy and I absolutely hated group activities, but I have found that I really like doing group activities in this course, I have managed to make friends just by doing them.

I don't believe there should be any changes made, I think everything laid out still makes sense for what we want to accomplish, the time is flying by and I think everyone in our class is managing to meet the goals well.

B3 finds the contract fair, and helping her meet her goals. She also finds the contract helping her colleagues, seeing our class in a more communal light. While she is concerned about her ill father, and offers some details that constrain her ability to come to class every day, she doesn't seem anxious about the course, or resentful about the labor involved in it. The one issue she raises is the late/missed class policy, which many other students in the class mentioned, a detail we renegotiated in the next class session. Furthermore, unlike Inman and Powell's students, she doesn't seem concerned about grades as markers of status or progress. The affective dimensions I read in her reflection have more to do with her relationships with her colleagues and their feedback on her writing, which have little to do with status and more with mutually beneficial and compassionate relationships.

T3 (Latina) also offers typical reflections on our T3 contract that demonstrate in other ways how different our labor-based assessment ecology is from those in the hybrid contract studies. After wondering about the number of missed/late days in the contract, T3 explains:

The contract, in my eyes, is still fair. As said in the beginning of the quarter, there are no surprises. Students know what to do in order to receive full credit. Everything is laid out for us and that is an extremely fair advantage that will only provide success rather than failure.

I believe that the grading contract and the things we have done in the class so far reflect the course goals well. This is a very engaged class, we are constantly doing group work and editing and reflecting each others assignments which is what makes this class so enjoyable. In the beginning of the quarter we were all awkwardly silent towards each other and did not have much to say about the things we read or the papers we wrote. But now, we all look forward to having group discussions and getting to talk about what we read/wrote.

The course goals were very specific on what Asao wanted us to accomplish and we are almost done with the quarter and finished these goals without even noticing.

Similar to B3, what appears to characterize the grading ecology for T3 is not a need for grading, nor a nagging affective desire a grade, but enjoyable relationships with colleagues. Reflections like these reveal, in my opinion, students who are at ease with the course, even if they may say that the course is a lot of work. I argue that the reason students like B3 and T3 can be so at ease in our ecology, expressing no anxiety over grades, or the leveling effect, is that the teacher's judgments of

writing are not used in the ecology to determine course grades at any level. There are no contradictions of judgment, at least not in terms of calculating grades.

Perhaps the one concern above about making the course more difficult than expected by enforcing more labor than students were used to (Concern #5) might still be an issue in my version of labor-based grading contract ecologies. But it looks different in my classrooms, and I think it can be seen as positive. Spidell and Thelin offer Rachel's response to the contract as indicative of this concern, in which she says the contract makes some students "angrier" because: "If they planned on showing up, doing minimum work and passing with a C, than [*sic*] forcing them to sign a contract that says that's what they have to get is a bit crazy" (49). While Spidell and Thelin identify this comment as evidence of Concern #5, I also see it as evidence of #2. My students rarely voice such concerns, but sometimes they do. However, because I ask them explicitly about the contract's level of fairness, I believe this helps students separate the amount to work from other concerns that I think are at the heart of the "crazy" concern here, and it leads back to the contradictions in judgment that are a part of hybrid grading contract ecologies.

But I have other evidence of such conclusions. Compare Spidell and Thelin's student responses to the generally high effectiveness, happiness, and preference rates of students at CSU, Fresno for labor-based grading contracts in two consecutive academic years (2009 and 2010). Considering just the White racial formation in these data, since they may more closely match the student populations that Spidell and Thelin use, we see some evidence of what they find, but with a key difference. At CSU, Fresno, we found that 68.32% (spring 2009) and 58.24% (spring 2010) of White students preferred labor-based grading contracts, and 74.75% (spring 2009) and 76.67% (spring 2010) found them effective, while 70.30% (spring 2009) and 70.97% (spring 2010) were happy with the contracts (Inoue, "Grading Contracts" 88).

While these students had mixed feelings about the contract in open-ended comments (with 117 positive and 109 negative), there were two strong positive themes in them, concerning "'relieves pressure' (29 comments) and 'expectations and clarity' (21 comments)" (91). Labor-based grading contracts offered students relief from the pressures of grades on their drafts and provided clear expectations for course grades. These end of semester survey responses seem to match my own students' reflections, like T3 and B3 above. The other racial formations at Fresno had similar comments, with fewer mixed reactions. In fact, the Asian student population, which were mostly Hmong in our sample (who were mostly multilingual and first generation students), tended to find the labor-based contract system one that offered them "freedom to write" and "tended to express appreciation and praise for the contract's ability to keep grades off

their writing and assignments” (89). Meanwhile their preference rates were also higher (81.48% in 2009 and 70.64% in 2010) (88).

These kinds of programmatic findings suggest, at least for many students of color, that labor-based grading contracts do not cause the concerns that Spidell and Thelin see in their mostly White, Midwestern students. But the bigger difference, I believe, is not the racial subject positions of the students, but the presence of the contradiction in judgment in the ecology that affects various groups differently, but still negatively. More recently in my own classrooms, if Table 7.8 is any indicator, even White racial formations find success in properly designed labor-based grading contract ecologies.

The second criticism that could be made about labor-based grading contracts and the way I’ve argued their effectiveness in this chapter, centers on discussions of “audit culture” and the regulation of labor in writing courses made by Mary Soliday and Jennifer Seibel Trainor.⁶³ Drawing on student interviews of juniors, Soliday and Trainor wanted to understand how students experienced their writing intensive courses. They found that the regulation of writing tasks tended to create mixed feelings in students (half the students experienced regulation negatively, and half positively) around a lack of freedom to exercise agency and make decisions about their writing (134). Taken from their student interviews, the authors explain regulation: “they [students] felt regulated when confronted with assignments dominated by mandates and rules, lists of ‘do’s and don’ts,’ required steps, and rubrics that ‘must be’ followed in order to meet assignments” (126). They link notions of regulation by students in their study to the anthropological concept of “audit technologies,” which “are not ‘neutral’ ‘practices’ but ‘instruments for new forms of governance and power’” (127). They do not deny the need for certain kinds of regulation in schools, but feel that “regulation has become too powerful” (127).

Soliday and Trainor sum up some of their findings:

When students perceived that their writing was regulated, they tended to equate writing with assessment, ascribing rhetorical exigency to rules and steps mandated in prompts [. . .] affect[ing] how students negotiated complexity; how they related to their teachers and peers; and how they described their authorship. Conversely, when students saw writing in terms of doing and making, they ascribed rhetorical exigency to composing or to disciplinary content and purpose. These students moved beyond instrumental views of writing, imagining themselves as authors with professional expertise. (129)

63 I thank Jonathan Hunt for bringing this criticism to my attention.

Inman and Powell also discuss Soliday and Trainor's concerns about audit culture and regulation in their grading contract classrooms. They see evidence of audit culture in the ways that grading contracts tend to limit teachers' abilities to control students' behaviors through grades (45-46). Teachers in their study simply could not use grades to reward or punish, and this troubled them. This led to a felt problem for some teachers. They explain: "The course contract, in their initial experiences, blurred the lines of their [teachers'] roles in passing judgment on student writing, on their identities as judges and juries of writing. And there were concerns that they, as the people who sanctioned these students, would be judged as ineffective teachers by the institution" (47). This anxiety over a lack of regulation by teachers, I think, speaks more to the Whiteness *habitus* those teachers embody, than a problem with assessment ecologies that have no grades in them. As I discussed earlier in this chapter, a White *habitus* often brings with it a set of dispositions toward being an authority and judge over others, and a sense of benevolence about one's judgments. I hear this *habitus* in Inman and Powell's teachers' desires to grade.

However, with my use of labor logs and labor instructions that are primarily process instructions, or step-by-step processes that lead students through reading and writing labors of the course, it is easy to see how regulation is a part of my ecologies. The question that I think Soliday and Trainor want us to ask is, what kind of regulation is occurring and what effects does it have on students' agency as writers?

It is important to first note that Soliday and Trainor's study is of juniors in writing intensive courses, not writing courses taught by trained compositionists. While I'm sure there are good teachers of writing in other disciplines, their training is limited, and they may not know or be able to apply many of the best practices that the field of writing studies holds. They also define agency in part by how closely a student takes on the *habitus* of one's discipline, which is an unrealistic expectation to place on FYW students, most of whom may not know what they want to do yet, or may not have picked a major. Agency in writing decisions for FYW students likely should be defined differently.

Despite these caveats, two things seem clear about the kind of regulation or audit technologies to which Soliday and Trainor refer and that have negative effects on student agency in writing. The kind of regulation set up in these negative assessment ecologies sends messages to students that writing is an assessment, a test of knowledge or ability, and that writing has mandated rules and steps to follow, which likely will be graded. So the nature of the regulation is to control the nature of what students produce. The audit technology is a check-list that allows the teacher to account for information learned or tasks completed. It is lists of to-do's, which include items in a draft that are listed on rubrics and prompts

(to be graded). Students know what this kind of regulation is attempting to do: control the products or outcomes of their learning. But as my discussion of why I based effectiveness on noncognitive goals shows, the audit technologies in my courses do not regulate the products of students' labors. It regulates their labors, and allows them to determine what those labors create and even mean to them.

In Soliday and Trainor's study's conclusions, they describe the more agentic students, the ones that they say take on writing as "craft" by experiencing it as a "subtle relationship between skill and creativity, rules and imagination" (129). Furthermore, this craft model also tended to be encouraged in courses that set up writing as a "sketch design," or as "a 'kind of bounded openness'" that encourages students "to connect technical craft to the imagination," meaning in most cases writing was not equated to assessment, yet procedures were still given clearly (136). Writing tasks, the labor of writing, then, in their study, was articulated by students as "doing and making." I'm not sure I could explain the heart of labor-based contract gradings any better than to say that it frames reading and writing practices as "doing and making." It keeps them as verbs, not nouns. And vital to doing this is making sure that writing is not equated to assessment. It is not an opportunity to be tested, but a chance to do something and make something. If most of this book has not made this argument in detail, then I don't think I'm able to make it.