Ungrading as Institutional Ecology:
How (Alternative) Assessment Choices Shape Writing Classrooms

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Megan Von Bergen
August 2023

Acknowledgements

I'm writing this acknowledgements section while listening to Of Monsters and Men, looking out my southern window watching the sun set across the trees. Somewhere beyond those trees are the Appalachian Mountains. Over the last six or eight months of my PhD, this is how I wrote most days — from 5-8 pm, with a drink, and with "King and Lionheart," "Your Bones," or "Little Talks" on repeat. I will miss these moments of concentration and peace, after I turn in this project sometime in the next few days.

Until then, however, I want to express my gratitude to the people and places who helped me complete this work.

First, to my academic colleagues who provided assistance as I worked on this project, thank you:

Josh Eyler, for looking over a version of the history of grades/grading in my archival chapter (Chapter 2). I appreciated knowing that section was pointed in the right direction.

Amber Roessner, thank you for such open-ended assignments in your Methods course. I used one of those Methods to return to the UTK Special Collections, to dig up an old composition syllabus I'd found there in Fall 2019 — and so, the archival chapter (Chapter 2) was born.

John Gallagher, for sharing your bibliography of case study research, and for tweeting about it. Our exchange (which came when half my dissertation was already completed!) helped me realize that I was writing a case study in the first place.

Seth Kahn, for a great discussion on Inoue's definition of labor/cultural capitalism. Your insights helped me work through how I envision assessment ecologies — and the importance of a more institutionally grounded approach, which I argue for in my Introduction.

Raul Pacheco-Vega, for your public dissertation-writing resources, especially the Dissertation Analytic Table. When I was stuck thinking about what, exactly, my dissertation argued and how the parts connected, that template helped me join up the dots and make a coherent argument.

My online communities, especially on Twitter. I suspect that some of the key interventions I make in this project — including, though not limited to, pointing out the overemphasis of ungrading on students, pointing out the conventional treatment of ungrading as an intervention – – come from discussions on Twitter, now lost to time. I wish I could cite the specific tweets or blogs or articles that proved meaningful to me, but those are long since lost to time. I want to express my gratitude here for the way that, particularly during two years of COVID and limited conferencing, those online conversations shaped my thinking and gave me confidence that my ideas would be taken seriously.

Laura Robinson, Matt Dowell, Paul Corrigan, Niki Hancock, and others for listening to me vent about the difficulty of coding and grad school relationships. You got me through some hard nights.

Allison McKittrick and Special Collections librarians, for their aid in finding (and re-finding when I lost them) the archival materials included in this dissertation.

Sharon Mitchler and Sonja Andrus, for your early support and encouragement during the entirety of my PhD program.

Lacey Wootton, Gin Schwarz, and Antonio Byrd, for sharing your own dissertations with me. Seeing what this genre turns out looking like was enormously helpful.

Lacey Wootton and Stacy Wittstock for talking about institutional ethnography with me.

And the most important for last, our excellent English Department administrative staff — Leanne, Judith, Pam, and Teresa. Y'all are the best, and I would not have finished without you.

Next up!

I'm grateful for the activities I took on which kept me grounded, especially ultrarunning. I ran more than 4700 miles between defending my prospectus in early June 2021 and submitting my dissertation almost two years later, in May 2023. I am grateful for the rest and peace that running, and running, and then running some more — especially on trails, through the southern Appalachian mountains — provided during a stressful time.

In that same vein, I am grateful for Tuesday night run club with the Long Run, for a once-weekly chance to get out of my own head by bombing around the UT Knoxville track as hard as possible. Thank you to all the friends I met there, who assured me there was more to life than dissertating.

And I'm thankful for No Business 100, which I completed in October 2022. That race — I thought I was insane to sign up for a hundred-mile race the last full year of my PhD, and then it turned out to be the best thing I did all year — taught me that I was strong and capable of mighty things, and taught me the only way to complete them — one loop, one step, one leg of the journey at a time. I brought that same spirit back to my dissertation, and completed it too.

I've been privileged to have good friends and family along for the ride as I worked on my dissertation, and I want to thank them as well:

Teresa, for your willingness to listen, your encouragement, and your counsel over the last few years. I also appreciate your keen input on an early draft of my theory of hostile/hospitable programmatic architecture, and I draw from our conversations in this document.

Jordan, for introducing me to key philosophical concepts and texts that shaped how I think about assessment, especially *Seeing Like A State.* Along the same lines, I'm thankful for some of our early conversations in Summer 2021 about assessment. Anytime you want to write that paper, I'm game! And I'm especially grateful for your decision to walk with me, and provide advice, when I was falling apart.

Kelly, we joke that you're my honorary committee member. It's funny because it's true. I appreciate you making space for me and reassuring me during those early, anxious months of the project, and your willingness to listen and talk shop — about doing grad school, ungrading, cats, and running. Thank you for crewing me through No Business 100.

My grandparents, who supported me financially during the five years it took to complete this degree, providing a little extra stability and assurance during several difficult years.

My parents, for your willingness to listen and your unfailing support of and belief in me.

And finally, thank you to:

My study participants. I learned so much from you. Your experiences with ungrading challenged me to grow and improve; I relied on your experiences in my teaching as well as in my research, and I would not be half the scholar I am if I didn't have the chance to learn from, and listen, to you.

My committee. I'm grateful for your thorough feedback on my writing; your willingness to talk through what I was writing and how I was feeling about writing it; your coaching on some of the difficult bits; and your support and advocacy of my work in the department, especially the Equitable Assessment Pilot (2021-2022) and, though they're not part of this project, the Student Success Grant (2022-2023) and the change to our comprehensive exam processes (2020-2021).

To Jessi, thank you for teaching me to use tables and graphs to make my points and to think quantitatively. Thank you for your early support of my emerging interests in antiracist writing assessment and education. To Lisa, thank you for working with me to change our comprehensive exam processes — an institutional change I remain quite proud of, and which convinced me that the UTK English Department would be open to further change; thank you for all your suggested revisions, which made this work most accurate. To Jud, thank you for asking questions which pushed me to think in new ways — always essential to meaningful work. Jeff, I am especially grateful for your willingness to meet me where I was and take on some unconventional advising responsibilities over the past few months. Thank you also for your insistence that what I was writing was good, even when I didn't believe it myself.

And my readers. May reading this project be as valuable, informative, and transformative for you, as it was for me.

Abstract

A longitudinal case study of grades and grading in the 1920s and 30s and the turn towards ungrading (2020-2022) in the First-Year Composition (FYC) program at the University of Tennessee, Knoxville (UTK), this project argues that institutional architecture structures classroom writing assessment and that the outcomes of ungrading (an umbrella term for a range of alternative assessment practices, including labor-based grading) vary based on teachers' values/beliefs about writing. While rhetoric and composition scholarship on writing assessment typically frames ungrading as an individual, classroom-level choice that improves learning and increases equity, this project approaches ungrading from an institutional perspective, focusing on how programmatic and university contexts shape the function of conventional and alternative writing assessment and teachers' experiences with ungrading.

Drawing on archival data from the University Special Collections, the project opens by arguing that grades/grading prioritize institutional needs/reputation over student learning, mandating the use of standardized American English. Analyzing gradebooks kept by English professor John C. Hodges (1926-1938) shows that grades assigned in first-year writing courses fall along a bell curve, artificially depressing students' grades and constructing students as in need of remediation. Grades do not track learning but rank students by the then-emerging standard of formal academic English. The project then jumps ahead a century to the contemporary First-Year Composition program (2020-2022), exploring the emergence of ungrading, or non-authoritative forms of writing assessment that center students' labor and experiences in the course. Drawing on qualitative data from interviews/focus groups with graduate instructors and non-tenure track faculty, the project shows that programmatic architecture is key in depressing or expanding the use of ungrading. A resistant or hostile programmatic architecture may cause instructors to limit their use of ungrading, but writing programs can provide a more hospitable institutional context by ensuring faculty have the permission and resources to use alternative assessment. When instructors do use ungrading, they experience its outcomes as variable, dependent on their own values/beliefs about writing. This variability also means that the longer faculty use ungrading, the more likely they are to see meaningful results from its use.

Table of Contents

Introduction 1

A short history of ungrading scholarship 8
Early efforts to ungrade 8
Emerging composition scholarship on ungrading 9
Ungrading scholarship moves into the mainstream. 10
Ungrading since 2020 11
Key Themes in Ungrading Literature 12
Assessment Ecologies are Rhetorical Ecologies 17
Rhetorical Assessment Ecologies — So What? 20
So where do we need to go next with ungrading? 22
Conclusion 27

Chapter 1: Methods of Inquiry: Case Study and Institutional Ethnography 29

Why focus on the University of Tennessee, Knoxville? 30
How does case study as a method(ology) affect my choices in data collection and analysis? 32
Generalizability 32
Locatedness of research and knowledge 35
Relationships 37
Multiple Methods 38
Use of institutional ethnography 39
Limitations to this case study 40
Institutional Limitations: Distinctive Features of the UTK English Department, 2020-2022 40
Individual Limitations 42
Positionality 43
Conclusion 49

Chapter 2: A History of Assessment in First-Year Composition at UTK, 1926-1938 51

A note on framing 53
How did *grading* come to be? 57
How did first-year composition come to be? 59
Purpose of first-year writing 64
Report on grade distributions 68
Results 71
Analysis 77
Qualitative data 82
Discussion 84
Conclusion 87

Chapter 3: Interlude 89

Institutional and social contexts 90
AY 2020-2021: Informal growth of ungrading 90
AY 2021-2022: Piloting ungrading 91
What do UTK instructors mean by ungrading (2020-2021)? 92

What do UTK instructors mean by ungrading (2021-2022)? 96
How do instructors at UTK experience ungrading? 99
Conclusion 101

Chapter 4: The Role of Hostile vs. Hospitable Programmatic Architecture in Supporting Ungrading 103

2020-2021 data 109
Boss Texts 109
Work Processes 112
Relationships 114
Conclusion, 2020-2021 117
2021-2022 data 118
Permission 119
Limitations to the permissory nature of the Equitable Assessment Pilot 124
Instructor Use of Administrative Resources 126
Instructor Departure from Administrative Resources and Guidance 128
Conclusion 131

Chapter 5: The Impact of Teachers' Values on Perceived Affordances of Ungrading 133

Methods 136
Covid and ungrading 137
Correspondence between Ungrading and Teacher Values 142
Illustrating the Convergence between Values and Ungrading Affordances — Three Examples 145
Carter 145
Sterling 152
Taylor 157
Ungrading's Affordances During 2021-2022 162
Sterling (2020-2021) 163
Taylor (2021-2022) 167
Conclusion 171

Conclusion 173

Writing Educators 177
Writing Program Administrators 178
Writing Researchers 180

References 182

Appendices 191

Appendix A: 2020-2021 Consent Form 191
Appendix B: Interview Protocol 2020-2021 195
Appendix C: 2021-2022 consent form 201
Appendix D: 2021-2022 Recruitment 205
Appendix E: 2021-2022 Interview Protocols 206
Appendix F: Collection and Audit of Archival Data 213
Appendix G: Grade Distributions, 1928-1929 through 1936-1937 215

Appendix H: Codebook 221
Vita 224

List of Tables

Table 2.1: Course Descriptions for ENGL 111, 112, and 113 in the University of Tennessee Catalog (1926-1927) 66

Table 2.2 Grade distributions, ENGL 111, 112, and 113, AY 1926-1927 72

Table 2.3 Grade distributions, AY 1927-1928 74

Table 2.4 Grades received by individual students in ENGL 111, 112, and 113, AYs 1926-1927 through 1936-1937 76

Table 2.5 Grade distribution of all ENGL 111 sections, AY 1935-1936, College of Liberal Arts Report, 1936-1937 81

Table 2.6 Qualitative feedback in gradebooks kept by John C. Hodges, 1926-1929 83

Table 3.1 Participants' assessment practices, as components of ungrading, 2020-2021 94

Table 3.2 Participants' assessment practices as components of ungrading, 2021-2022 97

Table 4.1 Codes With Definitions for Educator-Writing Program Interactions, 2021-2022 120

Table 5.1 List of teachers' values and beliefs about writing/writing education, Fall 2020-Spring 2022 138

Table 5.2 List of teachers' perceptions of affordances of ungrading, Fall 2020-Spring 2022 ... 139

Table 5.3 Instructor mentions of Covid as motivating their adoption of ungrading 140

Table 5.4 No. of data segments coded at particular values versus at ungrading affordances 143

Table 5.5 Values, listed by prevalence — Carter, 2020-2021 146

Table 5.6 Perceived affordances of ungrading, listed by prevalence — Carter, 2020-2021 148

Table 5.7 Values, listed by prevalence — Sterling, 2020-2021 154

Table 5.8 Affordances, listed by prevalence — Sterling, 2020-2021 154

Table 5.9 Values, listed by prevalence — Taylor, 2020-2021 159

Table 5.10 Perceived affordances of ungrading, listed by prevalence — Taylor, 2020-2021 ... 159

Table 5.11 Change in teacherly values, listed by frequency — Sterling, 2020-2021 to 2021-2022 164

Table 5.12 Change in perceived affordances of ungrading, listed by frequency — Sterling, 2020-2021 to 2021-2022 164

Table 5.13 Change in teacherly values, listed by frequency, Taylor, 2020-2021 to 2021-2022 168

Table 5.14 Change in perceived affordances of ungrading, listed by frequency, Taylor, 2020-2021 to 2021-2022 168

Table G.1 Grade distributions, AY 1928 – 1929 216

Table G.2 Grade distributions, AY 1929-1930 216

Table G.3 Grade distributions, AY 1930 – 1931 217

Table G.4 Grade distributions, AY 1931 – 1932 217
Table G.5 Grade distributions, AY 1932-1933 218
Table G.6 Grade distributions, AY 1933-1934 218
Table G.7 Grade distributions, AY 1934 – 1935 219
Table G.8 Grade distributions, AY 1935 – 1936 219
Table G.9 Grade distributions, AY 1936-1937 220

List of Figures

Figure 2.1 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1926-1927 72

Figure 2.2 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1927-1928 74

Figure 2.3 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1927-1928 through AY 1936-1937 75

Introduction

Amid Thanksgiving 2019, a year and a half into my PhD program, driving across country to visit family, I caught an interview with Black activist Andre Henry on a podcast. As I listened, Henry explained that those of us who live in a system constructed on systemic injustice do not have a choice to remain neutral, and be innocent. We can either identify with the marginalized and oppressed — or we can participate in injustice. Justice work starts, Henry said, with the belief that *It doesn't have to be this way* — a willingness to acknowledge the systemic injustice and harm in the current structure, and imagine other ways of being, partnering with other people across the communities we inhabit to work towards these more hopeful realities. This call came to haunt, and define, my teaching, service, and research across the next five years of PhD program — but especially in beginning 2020, as the pandemic and the resurgence of the Black Lives Matter movement threw a spotlight onto systemic injustice in our own society.

At the time, I was familiar with ungrading — specifically, labor-based approaches, following Inoue (2018) — and had already experimented with a hybrid approach in my first-year writing courses. But, assigned to teach a five-week summer course that year, I leaned into ungrading fully, shifting to an entirely labor-based approach in hopes that a more generous, transparent form of assessment could mitigate some of the chaos students were experiencing and better acknowledge the very real injustices they faced. I also began sharing about ungrading more broadly among members of my department — putting together an online folder with resources, meeting up with colleagues at a park to talk ungrading and answer common questions, giving talks about ungrading and antiracism, based on what I was finding in the literature.

I also designed a study to talk with colleagues at UTK about their experiences ungrading — what their motivations were for taking up the practice and what it was like using it in the classroom, during those chaotic months in the middle and end of 2020. I found that their reflection mirrored my own, seeing in ungrading a (partial) response to the ongoing injustice — including, importantly, the injustices and inhumanity perpetuated in the institutional spaces we inhabited together — of that summer and fall.

In this interview excerpt, Kerry and Lane (pseudonyms for two graduate student teachers who participated in that study) reflect as part of a focus group on their experience using ungrading in 2020. Speaking to each other, they say:

> Kerry: Teaching was the highlight of my summer and most of the rest of it was awful.
>
> Lane: Yeah. It really was this time of [pause] I - I felt so disconnected in terms of like, Here we are in academia discussing comparative rhetorical analyses and it's like people are dying, and people are being killed, and I don't — It just was this moment where it really brought to light like the privilege we have? And I'm like, and so why would I try to then give you a 92? Yes, I want you to learn and grow and if this can be kind of a mental escape, great! [pause] but, I don't know, it was just a lot of feelings.

> Kerry: Yeah, and I'll say that's something that my Twitter timeline was like full up of from like March till I mean now, really, of just people being like these are unprecedented times, we can't keep pretending that it's normal, and we're like going on as usual. This is an important moment to you know, either acknowledge that your students are human and have things that affect their life outside of the classroom that you take seriously, or to just steamroll them and ignore their suffering in the name of academic excellence. And I was like, Oh yeah! There's two clear camps on Twitter, and I want to be with the non-shitty humans right now."

Two things stand out to me about this passage, even nearly three years on, as I write this in Spring 2023. One is educators' clear commitment to humanity and well-being of their students – – and, importantly, of themselves as well. Kerry and Lane express a desire for students to have a "mental escape," to be acknowledged as "human," with "things that affect their life outside the classroom," and affirm the value and satisfaction they got out of teaching that summer, in their ungraded courses. But the second point, also important, is the tension between educators' desire to affirm and support their students', and their own, humanity and the goals of the larger institutional *habitus* where they dwell. Were academia to "go on as usual," as Kerry explains, its centering of academic excellence — its insistence in the primacy of its own objectives and goals, the importance of academic assignments, such as the comparative rhetorical analysis, that may or may not bear any relationship to the ongoing difficulties of real life — would wind up "steamroll[ing]" students. Educators are offered a choice, between participating in and maintaining the typical norms and structures of higher education and being "non-shitty humans." Higher education, as Kerry and Lane discover, resists the drive to make another world.

I take up this tension in this project, exploring how academic institutions assert certain values, among them perceived intellectual rigor and exclusivity, through grades and grading in first-year writing courses — and then, how academic institutions may be remade along more hopeful lines, dismantling the structure(s) that resist the turn towards "non-shitty" ways of grading and making space for liberatory assessment practices. Along the way, I discuss the motivations that writing educators bring into ungrading and how those motivations affect the affordances or outcomes of ungrading within the classroom, making it a dynamic, fluid element of the wider assessment ecology — meaning, the (institutional) spaces or environments where assessment work happens, from the classroom to the larger department and university. I do this work via case study, focusing on a single institution: the University of Tennessee, Knoxville's First-Year Composition Program. I trace the history of grades in first-year composition at the University of Tennessee, Knoxville and then (amid the upheaval of the years following 2020) the turn towards ungrading among contingent faculty (GTAs and NTT teachers) in the Composition Program here, telling stories that educators, including Kerry and Lane, share with me in hopes that writing educators and administrators at other institutions may find this work instructive in imagining how the forms of assessment they use may be different — and more humane.

My goal is that this work may inform and guide decisions other institutions make, the way they pull down, rebuild, and renovate the programmatic and departmental *habitus* where students, teachers, and administrators dwell and work together in ways that are alive to the ethical importance of assessment practices that do not rank students against a single standard (as was

historically common) but in fact leave room for educators to employ and develop pedagogically sound, antiracist forms of writing assessment — writing assessment that seeks to not (only) enculturate students into academic writing expectations but also affirms their own language(s) as rhetorically powerful (see Paris pg. 95, 2012[1], see also Baker-Bell 2020). I want too for this work to exemplify the hope that brought me to it, that *It doesn't have to be this way.* By charting out how the First-Year Composition Program at the University of Tennessee, Knoxville, changed its writing assessment practices, I illustrate that such change is possible, laying a course for future work by writing educators and administrators, calling them, too, to make educational choices that work for justice.

In the next few pages, the introduction to this project, I trace how I come to this focus on academic institutions and the writing educators who inhabit them through the prior literature on ungrading. Much of the work so far on alternative assessment in composition courses centers, as I show, the effects of the grading process itself. Such work argues that particular models of alternative assessment such as grading contracts — which typically shift attention away from qualitative judgments of student works towards their fulfillment of certain, clearly-articulated requirements — are good for or interfere with learning, or that they work for or undermine equity among students[2]. This is important and valuable scholarship, pointing readers towards the mostly positive and affirming impacts of alternative forms of assessment in teaching writing. Yet as I show, this work is also limited, documenting primarily the effects of ungrading, as a practice, without engaging the substantive, important questions of how teacherly experience and motivation affects outcomes or, more broadly, how the wider institutional architecture where teachers work — and which their work, in part, constitutes — constrains or invites liberatory forms of assessment. This inattention is, I argue here, at least in part due to the fact that scholarship on alternative assessment, and assessment ecologies, does not engage with *rhetorical* ecologies, which provide a more robust theory of intertwined, overlapping, mutually constitutive ecologies that speaks to how both writing educators' motivations and programmatic architecture may have unintended effects on ungrading in writing programs. I conclude the introduction with a road map for my own project, linking my use of case study and institutional ethnography (LaFrance 2019) as methodological frames for my study with my careful attention to the relationships between programmatic structure and teacherly assessment decisions. As my road map, or outline for the project shows, I address the gaps in literature on ungrading in first-year

[1] Django Paris urges educators to take up "culturally sustaining" pedagogies, meaning pedagogies that "support young people in sustaining the cultural and linguistic competence of their communities while simultaneously offering access to dominant cultural competence" (pg. 95)

[2] Surprisingly, few or no robust definitions of "grading contracts" exists, even in the most dominant literature on the subject (Danielewicz and Elbow 2008, Shor 2009, Inoue 2019). While, typically, grading contracts deemphasize teachers' qualitative judgment of how "good" student work is, assigning grades based in whole (Inoue) or part (Danielewicz and Elbow) on their completion of certain tasks, this is not universally true; and as Michelle Cowan (2020) points out, some versions of contracts continue to evaluate students' work based on (a) qualitative standard(s). (See Shor for an example of the persistence of qualitative judgment in a contract graded course.) Some scholars define contracts as an "agreement" (Medina & Walker, 2018, p. 51) with students, though unilateral approaches, including those used by Danielewicz and Elbow, do not fit with this definition. In my own courses, I use an entirely labor-based approach — I provide students qualitative feedback, but that feedback is not attached to the grade students receive. As I tell my students, "If you do the things, you'll get an A." Additionally, some contracts may be negotiated with students (Shor, Inoue), while others, including my own, are unilateral.

composition courses, inviting writing educators and administrators to grapple with the reality of the ways our programs support "shitty" ways of grading, oriented towards rigor and standardization, and making space for both institutional activity and individual teacherly choices to develop dynamic ways to think about and adopt liberatory forms of classroom assessment.

Before I turn to the scholarly literature, I want to take a moment to be clear about how I am using the term, *ungrading.* Later in this project — in the Interlude — I provide an emic definition of ungrading, using writing educators' own practices and affective experiences to construct a definition specific to the University context and growth of alternative assessment from 2020-2022. Here, I want to clarify my definition of *ungrading*, often, especially in cross-disciplinary pedagogical conversations, associated with a specific version of alternative assessment: students determining their own grade in consultation with instructors (Eyler 2022). However, I deploy it, within this conversation, grounded in rhetoric and composition, as an umbrella term, to capture a range of alternatives (sometimes used individually, sometimes used in combination with each other) to the conventional practice of awarding points (such as 9/10), percentages (92%) or letter grades (A) to student work based on the perceived quality of that work. Having an umbrella term is key. In rhetoric and composition, the most common version of alternative assessment is contract grading, which has an extensive history (Danielewicz and Elbow 2008, Shor 2009, Cowan 2020) but is currently prominent thanks to Asao Inoue's strong advocacy for labor-based contracts (2018). Using "labor-based contract grading" as my central term for alternative assessment practices, however, implies a certain set of practices, arranged in certain ways, glossing over the ways that alternative assessment practices are often remixed and rearranged to meet writing teachers' pedagogical goals (Schwarz 2020)[3]. Even within rhetoric and composition, the range of possible assessment practices, from collaboratively developing assessment criteria with students (Inoue 2004) to asking students to log their labor (Inoue 2018) to qualitatively grading student work (Danielewicz and Elbow 2008, Shor 2009), varies widely, making "labor-based contract grading" an overly precise term for what I am interested in, the move away from qualitative, authoritative judgments of student work towards more expansive alternatives that open up success to as many students as possible (Elliot, 2016). Thus, while *ungrading* in the context of rhetoric and composition usually signifies some version of labor-based contract grading, it does not *only* signify that, and in the pages that follow I use it to capture, within writing education and writing courses, a broad, fluid set of assessment practices that, by providing an alternative to points-and-percentage based grading, seek to affirm student learning and improve equity.

I also want to be clear that while I frame ungrading as a tool to improve equity, it is a potentially or partially antiracist project, not an inherently antiracist one. A robustly antiracist approach to ungrading is possible — see my discussion of Sterling in Chapter 5 — but not automatic, requiring years-long, purposeful attention to the connections between who students are, and how and why they use language and rhetoric. *Antiracism,* as framed in composition and literacy

[3] For an example of the problems with deploying "labor-based contract grading" or even "contract grading" as a primary term for a certain set of assessment practices, see Kryger and Zimmerman (2020) and Carillo (2021): the scholars provide essential critiques of the conflict between labor-based approaches and the needs of neurodivergent/disabled students, yet because they focus in part on labor logs as the locus of their critique, a practice which not everybody uses, their critique falls short.

education scholarship, typically has two components — first acknowledging and dismantling systems biased against minoritzed students; and second, critically engaging students' linguistic, rhetorical, or cultural practices as core to their identity. Dismantling biased or discriminatory systems is key to preventing harm and opening up space for more equitable practices, while purposeful engagement with the way students' racial identities affect their choices as writers and their academic/professional goals ensures more affirming education (see Sanchez and Branson, Paris, 2012; Pimental et al 2017, Randall et al, 2022). Antiracist teaching, though it offers students the autonomy to use standardized or "dominant" forms of language/culture, validates their culturally-informed writing practices as rhetorically powerful (Baker-Bell 2020, Paris, 2012).

This two-pronged definition, at once deconstructing biased systems and reconstructing more equitable, "culturally-sustaining" (Paris, 2012) ones, extends to assessment as well. Conventional forms of assessment are indeed historically biased against linguistic or rhetorical choices outside of those associated with whiteness. First-year writing historically enculturated students into the white language norms required for managerial positions (Berlin, 1987) — as I discuss in the archival chapter (Chapter 2) — and is still experienced as violent by students from communities other than white, middle-class ones (Yagelski, 1999). Black and brown students[4] are often assumed to require additional instruction in the "right" way to use language, and they are reminded by those in authority that their language is inappropriate for academic or workplace writing, in court, and in public protest and self-expression (Baker-Bell, 2020, Jordan, 1988)[5]. Antiracist assessment, then, acknowledges on the one hand the deep interconnection between whiteness and social ideals of publicly-acceptable, professional language and seeks to tear down the conventional grading practices that reinforce that connection, while also explicitly engaging the strength and capacity of students' own home languages, rhetorics, and cultural practices. Randall et al (2022) write:

> A justice-oriented antiracist assessment approach … require[s] an explicit confrontation of racism in our assessment practices and works to disrupt these systems of oppression through assessment practices specifically designed to sustain, not eradicate, students' cultures, languages, and ways of knowing/being. (Inoue, 2015). Antiracist assessment

[4] My use of the term "Black and brown" to describe racially minoritized students' follows the field's usage and reflects the fact that many of the scholars I cite in this project, from April Baker-Bell and Sherri Craig to Asao Inoue, identifies as Black and/or brown. However, I acknowledge that "Black and brown" erases Indigenous students' and teachers' experiences, and that BIPOC, or Black, Indigenous, and People of Color, is preferred. In future iterations of this project I hope to engage more fully with Indigenous scholars writing about assessment and liberatory versions of education and expand my terminology then.

[5] Jordan's essay describes a college class of Black students who, when one students' brother is killed by the police, write a collective letter of protest to local newspapers, hoping to honor the brother's memory and make their own outrage at this injustice widely known. But (as Jordan notes) the students have a choice to make: to write in standardized English, or to use their own Black English; to "use the language of the killer" to "make our ideas acceptable to those controlling the killer," or to maintain "the language of the victim, Reggie Jordan." The students ultimately choose to use Black English for their protest letter, capturing their grief in a powerful, eloquent message that reflects their own identities, and the identity of the victim. Yet — as the students and Jordan all knew — that choice prevents their letters from being published. Mass media prefers standardized forms of English.

> interrogates the ways in which racism rationalizes and continues to perpetuate injustice and support differential hierarchical power structures (based on race) in society.

I want to highlight two things about this definition. First is its attention to two sides, deconstructive and reconstructive, of antiracist assessment — at once disrupting oppressive or biased assessment practices (Randall et al specifically targets validation methods for large-scale testing) while also developing new practices, steeped in a critical awareness of power and race, that better sustain students' linguistic identities. This work is key. As April Baker-Bell argues, "[Educators] can't be out here using these mediocre and problematic measures of success that only legitimates a white status quo 'American Dream' . . . tethered to the death of blackness and Black Language" (30). Instead, antiracist assessment develops new ways to engage student work, honoring who students are and what goals they have for their writing.[6]

Also important is that by referencing students' "ways of knowing and being," Randall et al underscores that antiracist assessment responds affirmatively not only to diverse Englishes (linguistic diversity) but diverse rhetorical moves and ways of meaning-making in composition (rhetorical diversity). Conventional assessment assumes, and measures students' work against, writing structured and developed to reach an educated, white academic reader. But in fact, students may produce rhetorically excellent work that draws on more diverse genres, digital/material resources, evidences, or cultural references intended to be persuasive for readers from a minoritized group (Cedillo, 2017). Antiracist assessment drops narrow standards for what counts as good academic writing and instead validates students' diverse composing choices, linked to their own racial identity and rhetorical/academic goals.[7]

I want to be careful in applying this definition of antiracist assessment — at once identifying and deconstructing discriminatory practices; and cultivating more critical, culturally-sustaining ones in their stead — to the practice of ungrading in first-year composition at UTK. Ungrading does the first thing, for the most part; it does not necessarily do the second. UTK is a predominately white university (PWI), and most first-year composition educators are white, without a robust understanding of all that antiracist pedagogy entails. As they (we) take up ungrading, they (we) blunt the impact of grades/grading on minoritized students; but lack the knowledge or resources necessary to validate students' diverse languages and rhetorics. I discuss the impact of my own whiteness on this project in my methodology, Chapter 1. Here, I want to note that approaches to ungrading as an antiracist practice varied widely among the first-year composition instructors I spoke to for this project. Some engaged students in detailed, robust conversations about language and identity; others expressed reservations about whether they knew enough about antiracist theories of language/writing to facilitate such conversations; still others continued to rely on

---

[6] The actual assessment methods used to challenge discriminatory language and rhetorical standards encoded in conventional grades and grading practices vary. Versions of labor-based grading, as Inoue advocates for it (2018), are perhaps most common, though self- and collaborative assessment (Litterio 2018) also play a role in moving away from qualitative judgments of student work based on a white standard, while other educators, concerned about equity for disabled students, use specs grading, a version of pass/fail grading that asks students to demonstrate competency in certain clearly-established learning outcomes (Kryger and Zimmerman 2020, Nilson 2014).

[7] Many scholars have well-developed definitions of rhetorical diversity. I find Sheila Carter-Tod's blog post (2021), "Rhetoric(s): A Broader Definition," insightful, a good place to start.

practices associated with grades/grading, for instance assessment tools (such as but not limited to rubrics) that reaffirmed a narrow construct of what counts as good writing. My point here is not to critique my colleagues' use of ungrading or familiarity with antiracist approaches to composition pedagogy and assessment. My point simply is to stress that in our institutional context, ungrading as a form of antiracist assessment is messy and imperfect. Ungrading — however partial — at the very least interrupts the work of disciplining students into the singular, white discourses of academe; it pushes us as educators to reckon with the assumptions about good writing that drive our assessment choices. But ungrading, here at UTK, does not automatically or always critically engage with language and rhetoric, or explicitly affirm students' linguistic choices. Some educators do this work. Others do not. My framing of this project as inquiry into antiracist assessment, then, is hopeful and tentative — acknowledging that ungrading throws sand in the gears of historically discriminatory grading practices and invites us to do better, while also clarifying that the essential work of opening space to think critically about linguistic/rhetorical diversity is, at least here at UTK, is barely begun.

One more brief note. I use the term *equitable* assessment at times, instead of *antiracist* assessment. I tend to use this term in three ways. First, in discussing secondary literature in particular, I use "equity" to gloss discussion of the ways that alternative forms of assessment dismantle classroom hierarchies, as such discussions sometimes, but often do not, link those hierarchies to students' identities — especially their racial identity. Second and relatedly, using "equity," as a sort-of synonym for antiracist, allows me to gesture towards the ways that writing assessment does not only discriminate against Black and Brown students' language use but also against the ways of being and communicating that other minoritized groups claim — among them, disabled and neurodivergent students' composing practices (Carillo 2021, Kryger and Zimmerman 2020). Equity is intersectional work, and I hope this more expansive use underscores that though raciolinguistic diversity is, justifiably, a dominant theme in ungrading literature, given the history of writing/writing education of suppressing Black and Brown ways of using language, we as writing scholars and teachers are also obligated to work for assessment that engages and affirms the values of multiple ways of making meaning, including those linked to other facets of students' (and our own) identities (see also Carillo 2021). In fact, UTK educators I spoke to often framed their use of ungrading as allowing them to effectively support not only raciolinguistically diverse students but also students likely to struggle for other reasons (disabled/neurodivergent students, socioeconomically disadvantaged students, etc). The third and final way I use "equity" is in reference to the Equitable Assessment Pilot that the English Department sponsored (2021-2022). There, the term "Equitable" was in part chosen, quite simply, because the term "antiracist" was not safe in Tennessee[8]. The term "equitable" also invited teachers to look towards more expansive understandings of how ungrading could (re)shape their classroom to make it safe for all students — students of diverse socioeconomic backgrounds, genders, sexualities, neurotypes, and religious commitments. Ultimately, the more expansive phrasing of "equity" calls attention to the importance of assessment choices that recognize and account for the ways that conventional ranking of students' work against a

[8] Ironically, nearly two years after the Equitable Assessment Pilot got underway, I suspect that "equitable" is just as dangerous a word, given that right-wing reactionaries typically see it as code for work associated with critical race theory.

singular norm functions to bar folks from many minoritized populations from academic and professional success.

The next section in this introduction turns to a brief history of ungrading scholarship, tracing how (beginning in the 1960s and 70s) as the conversation grows, two key themes emerge: the potential for ungrading to cultivate a more pedagogically sound learning environment for students and increase equity, first by flattening student-teacher classroom hierarchies and then by addressing disparities in racialized judgments of language. With this historical overview as context, I then return to these two themes, showing how the twin emphases on learning and equity — both benefits that students experience, in the context of a single classroom — leaves open larger-scale questions about teacherly motivations for ungrading and the institutional structure(s) necessary to support alternative forms of assessment. I suggest that expanding Inoue's conceptualization of assessment ecologies (2015, 2019) to include theories of rhetorical ecologies (Edbauer 2011, Gries 2015) could cultivate an approach to assessment more keenly sensitive to the role of institutional structures in shaping and directing classroom assessment, and to the importance of teachers' own motivations and values on the actual classroom impacts of liberatory forms of assessment. This work, I conclude, could help address questions about the limits of ungrading scholarship to address or work for systemic change, as well as explaining the variation in the impacts of ungrading in classroom contexts; and invites writing administrators, teachers, and scholars, to situate their assessment work in its programmatic contexts and with an eye to their own values and goals.

A short history of ungrading scholarship

Ungrading scholarship has varied widely in terms of both quantity and focus over the years, sometimes flourishing as educators explore new ways of using assessment to promote learning and equity and sometimes dying away. I trace its history here, with attention to a few key themes: the emphasis on praxis over theory and the attention given to both pedagogy and equity as advantages for ungrading. For readers unfamiliar with ungrading literature, my hope is that this brief history provides a birds-eye view of where the conversation started, where it's been, and where it's going next, providing context onto which they can map my subsequent claims about the importance of attention to teachers' impact on the effects of ungrading and to programmatic or institutional supports for alternative assessment practices.

Early efforts to ungrade

Among the first appearances of ungrading in scholarship is the emergence of "de-grading" (Fleming et al 2011) at the University of Wisconsin in the late 1960s. "De-grading" specifically deployed pass/fail or other grading alternatives to boost students' GPA, preventing them from being drafted into the Vietnam War (p. 117)[9]. Additionally, TAs advocated for grading alternatives to correct for the inherent power imbalance between students and teachers in writing

[9] As Cowan (2020) shows, interest in the problems with grading and possible alternatives to it exist almost as long as grades themselves do — emerging in the early 1900s. But a lot of this work is not specific to composition, or composition at the collegiate level; the advocacy among TAs feels significant as it represents sustained, disciplinary interest in alternatives to conventional forms of grading.

classrooms (119). While these efforts did not always work, given that students struggled to take full use of the freedom given to them by assessment alternatives (Fleming et al 119-120, 2011), the advocacy for and sustained interest in de-grading and other alternatives to letter grades among TAs in the 1960s represents one of the first key moves towards alternative assessment in composition.

Emerging composition scholarship on ungrading

In the 1970s, ungrading appears as the subject of intense discussion in major composition journals, especially *College Composition and Communication.* Publications are consistently practical, not theoretical or empirical, meaning that for the most part, teachers simply recount their personal, pedagogical experiences using ungrading, often in response to their awareness of the problems with conventional grading and a desire to improve student learning. Teachers argue that grades fail to effectively communicate teacher feedback (Knapp, 1976), that grades distract students and interfere with their learning (Birdsall, 1979), and that grades fail to represent student work over the term (McDonald, 1975). In exchange, teachers propose a range of practices, from replacing grades on transcripts with narrative descriptions (McDonald, 1975) to delivering a pass/fail grade to students in one-on-one conferences (Knapp, 1976). A few texts report on program-wide use of alternative assessment practices, observing that though ungrading is more compassionate, educators find it to be insufficiently rigorous (MacDonald, 1975, at the University of Toledo) or arguing that a lack of clarity about the characteristics that constitute a passing paper produces curricular variation and unfairness (Purdy, 1975, at the University of Minnesota-Morris).

Much of this work, importantly, represents only a partial departure from conventional grading practices. Even teachers who argue against grades during the term may give grades — presumably, based on their own judgment of the quality of student work — by the end of term (see Birdsall, 1979 or Viet, 1979). Alternatives to grading often perpetuate conventional grading norms, such as qualitative judgment of student work, only this time without points or percentages signifying quality (Bucholz, 1979). Overall, then, though publications demonstrate interest in refiguring how grades are done to improve student learning, this remains a very small step in the direction of improving writing education and assessment, leaving the larger project, of assigning a grade to student based (largely) on their linguistic performance, intact. Additionally, with the exception of pieces such as Purdy's, the framing of ungrading remains focused on individual teacher choices within their individual classrooms — a point that, forty years on, becomes significant, as I explain later in this chapter.

After the 1970s, however, interest in ungrading seems to drop off. Much of the work published on ungrading in the 1980s, 90s, and even the early 2000s appears outside of flagship journals, in other spaces focused on practice and pedagogy, such as *Teaching English in the Two-Year College* (see Carter and Rockson 1998, Helton and Sommers 2000, Reichert 2003). A few high-profile scholars continue publishing on assessment alternatives — Peter Elbow, for instance, works out his theories of the importance of responding to student work in non-evaluative, affirming ways during this time (1993, 1997a, 1997b) — but compared with the 1970s, attention to ungrading is slimmer. The clustering of attention to alternative forms of assessment in

teaching-focused journals suggests that for better or worse, research and writing on (un)grading is seen as belonging to scholarship on teaching and learning; it's not "real" scholarship[10]. Still, work from Elbow and others lays the groundwork for the subsequent turn towards ungrading in higher-profile scholarship, beginning in the mid-2000s.

Ungrading scholarship moves into the mainstream.

Ungrading research, starting in the mid-2000s, moved back into mainstream composition journals, with scholarship making one of two claims (or sometimes, both): that ungrading improves student learning, or that it makes for a less authoritative, more equitable classroom. To the first, scholars argue that ungrading allows students the mental headspace or lack of anxiety around grades to focus on writing (Danielewicz and Elbow 2008, Shor 2009) and through cultivating the habits of mind, such as self-assessment and metacognition, necessary to being a good writer (Inoue 2004). An echo of earlier work from the 1970s, this scholarship leans into and hones the argument that ungrading improves students' learning in classes, focusing on specific characteristics such as metacognition, agency/autonomy, and clear objectives that are consistent across many (though not necessarily all) grading models. Not all the research centering student learning agrees that ungrading improves the classroom experience, however. Some writers suggest that ungrading models, especially versions of contract grading, are perceived by students as confusing and unpredictable, making learning less pleasant (Spidell and Thelin 2012). Others suggest that both students and teachers are so thoroughly enculturated into grades that operating without them as either a tool for motivation or a guide to one's standing in the course is difficult (Inman and Powell 2018). By and large, though, researchers readily frame ungrading as a benefit to student learning.

Other research emphasizes the potential of ungrading to work for equity. By giving students a voice in negotiated contracts, by making the criteria on which they are graded more transparent and accessible, alternative forms of assessment (partially) dismantle the innate hierarchical structures of the university classroom (Shor) and/or to establish standards by which they are, themselves, judged (Inoue 2004). Attention to equity, however, remains largely focused on student-teacher power structures, glossing over the role of identity — especially race, but also dis/ability, gender, and sexuality — in shaping students' classroom experience.

This changes with Asao Inoue's research, especially his two twin monographs: *Antiracist Writing Assessment Ecologies* (2015) and *Labor-Based Contract Grading* (2018). Here, Inoue argues that by moving away from assigning value to work based on qualitative judgments of students' performance of standardized, academic English, teachers can cultivate spaces that invite a wider range of values, or a more expansive construct of what "counts" as good writing, making it possible for students to consider how their own identity and home languages affect their rhetorical choices. The specifics of Inoue's argument I leave for later. What is significant here is that Inoue's scholarship represents a turning point in attention to ungrading as an

[10] This is not to say that there is *no* scholarship on grading alternatives in the flagship journals. Proportionally, however, far more research on grading alternatives exist in journals dedicated to pedagogy and teaching colleges than in research or theoretical journals. Between 1998 and 2014, for instance, *Teaching English in the Two-Year College* ran at least eight articles on contract grading; *College English*, for comparison, ran only one near that time period — one of Elbow's pieces, published in 1993.

equitable practice, from a focus on student-teacher relationships to more precise focus on students' identities as influencing both their power in the classroom and cultural milieu and the responsibility of assessment to account for who students are and how they use language. Especially in the years since, as Inoue's argument has become more prominent against the backdrop of rising anti-Black racism in the United States, the increased, explicit emphasis on assessment as antiracist is an important move.

Ungrading since 2020

Asao Inoue spoke at Cs 2019 about the role of labor-based grading in cultivating racial equity. Almost a year later, COVID lockdowns began, with protests for racial justice a few months after that — sparking renewed attention to labor-based grading and ungrading broadly in composition. This work has followed several directions. One particularly interesting direction is research exploring how antiracist assessment choices may be inequitable for students belonging to other minoritized identity groups — for instance, disabled students, as some scholars suggest that contract grading may assume a singular, easily-measureable way of writing that excludes students who are not neurotypical (Kryger and Zimmerman 2020, Carillo 2021). A second compelling direction is critiques of Inoue's thesis, arguing that in fact, ungrading does not address the needs of Black students (Craig 2021) or improve Black students' learning experiences (Carillo 2021). Scholars may call for equity work that is more systemic in nature, partnering with colleagues outside of the classroom to serve Black and brown students' needs (Barrios et al 2023). I address these critiques later in this chapter.

Much scholarship, however, remains diffuse, often echoing earlier themes in the literature, among them the potential for ungrading to improve students' learning experiences (Mallette and Hawks 2020, Stuckey et al 2020) or work towards antiracist ends (Tinoco et al 2020). Some of the work explores how ungrading works in contexts outside of a singular classroom, such as all-online programs (Stuckey et al 2020) distributed throughout a department (Tinoco et al 2020), or deployed in community colleges (Klotz and Whithaus 2021). Still other work considers how using ungrading may invite teachers to reconsider other assessment practices, such as feedback (Wood 2020) or participation (Gomes et al 2020), or how ungrading makes it possible to connect students with resources such as the Writing Center (Klotz and Reardon 2022). At least one dissertation shifts the focus from students' experience of ungrading to teachers', inviting readers to consider how educators put together their ungrading models (Schwarz 2020). All this work is good and important, especially for writing educators wanting to think through how ungrading may work in their own contexts. Yet I am struck by the fact that for the most part, this scholarship does not advance the conversation about ungrading, working from much the same premise as prior work in the field, especially Inoue's — the attention to ungrading as, primarily, a classroom endeavor undertaken to improve learning and equity outcomes. By focusing on "tweaks" that educators can make, such as adjusting participation credit, or applying ungrading in new contexts, this work leaves open questions about how and why writing educators come to ungrading and the specific supports necessary to sustain the use of liberatory forms of assessment at the programmatic level.

My own project takes up these questions, seeking to move the conversation beyond its current limits — the twin focuses on the way ungrading cultivates learning and racial equity, in

individual classrooms run by individual teachers; and the ways to tweak ungrading to more fully achieve these outcomes — by exploring the institutional architecture that supports, or dismantles, grades and ungrading, through the particular histories of the First-Year Composition Program in the University of Tennessee, Knoxville. Additionally, by exploring individual teachers' motivations within the programmatic context of First-Year Composition at UTK, I provide, in contrast with the largely individualized narratives of much recent work, a birds-eye view that allows me to explain how ungrading reaches varied, even divergent, ends. As a scholar-practitioner convinced of the value ungrading has for both racial equity and a robust, student-centered learning environment, I nevertheless hope that my institutional focus, combined with my attention to the experiences and commitments individual teachers bring to the work of ungrading, can extend the conversation about ungrading in ways that push past the established (potential) benefits of ungrading and possible tweaks to it to (re)consider how our own motivations and institutional contexts affect our work in this area. In the next section, I (re)turn to reviewing scholarly literature on ungrading, exploring the two most common themes, around learning and racial equity, more fully and showing how my dissertation picks up where the conversations leave off.

Key themes in ungrading literature

Two key themes in ungrading scholarship — its benefit to learning, and to equity — are of interest to me in this project. Of these two themes, the emphasis on ungrading as benefiting learning is the more common, as the short history above suggests. As I document above, writing teachers maintain a strong, persistent interest in the role ungrading — whether a grading contract, collaboratively written rubrics, or some other model — plays in learning. Research argues that ungrading hones students' metacognitive thinking (Inoue 2004, Litterio 2018), frees students to write what they want, without the pressure of evaluation (Danielewicz and Elbow 2008, Inman and Powell 2018), gives students more agency and choice over their work (Danielewicz and Elbow 2008, Inman and Powell 2018, Litterio 2018), invites a wider range of engagement from students, based on their learning preferences (Gomes et al 2020), helps students recognize that good writing is rhetorical — fluid and context-dependent (Litterio 2018), and motivates students by giving them control over the grade they earn in the course (Marchionda 2010). Overall, educators — and the students they teach, who in some cases (Gomes et al 2020, Mallette and Hawks 2020) coauthor these articles — clearly articulate a positive link between alternative forms of assessment that scale back or do away with grades and learning. Ungrading improves both the affective experience of learning, improving motivation and a sense of enjoyment in writing; and constructs the particular mental models of writing that rhetoricians hope to develop in their students, specifically sensitivity to the variable nature of good writing and an ability to make that decision for themselves.

This does not mean, of course, that every educator is on board. As I discuss above, some educators express concern that students' affective attachment to grades as a signifier of their own identity, relative to their writing, complicates the pedagogical value of ungrading (Inman and Powell 2018), or it suggests that students find ungrading confusing and unnecessarily opaque as a practice (Spidell and Thelin 2012). On the one hand, I want to be clear that I appreciate these critiques. We go awry if we eagerly accept ungrading as the solution to all ills in our classroom and implement it unthinkingly, expecting to transform students' learning experience with little

additional effort on our part. On the other, however, I want to suggest that these critiques are limited in several ways. First, as Albracht et al argue (2018), ungrading models vary widely and shape the impacts of ungrading on the writing classroom. Notably, of the scholarship critical of ungrading as a pedagogical tool, some pieces do not document the specific ungrading model used, as Albracht et al point out of Inman and Powell. Other pieces *do* share their ungrading model but choose one that is confusing and contradictory, as for instance Spidell and Thelin, who — drawing on Shor's model — continue to qualitatively grade student work along a single standard as part of their grading contract. My point here is less to critique either of these particular pieces than to suggest that determining the quality of ungrading relative to learning — does it help or hurt students' growth? — is a difficult and thorny question that cannot be conclusively answered outside of local context(s), since the model(s) of ungrading used vary so widely, as do teachers' goals with taking up ungrading.

Indeed, as I suggest later in this project, teacherly goals play a key role in determining the outcome(s) of ungrading. Much of the scholarship on ungrading focuses on its impact or effect on the course and/or on the students taking the course, much less on teachers' effects on ungrading. This relationship, however, shows up as key even in extant scholarship. In their discussion of the drawbacks of ungrading, Inman and Powell recount as a teacher — they take pains to tell us he is a "respected" (45), well-liked educator — who finds, when ungrading, that he misses the punitive element of conventional grading. He says: "the fear factor aspect of getting an F within the first month of your first semester of college, I think personally is sometimes enough, and I grade normally, in a normal class, grade really hard on the first one to let them know like there is something at the end here" (46). I don't want to make too much out of what one educator in one program said, but given that Inman and Powell reflects on the drawbacks of ungrading, I suggest that it is worth considering how teachers come to ungrading and what their motivations are in taking up the practice. Ungrading is deeply rooted in commitments to dismantling educational hierarchies, calling into question received wisdom about the way Things Are Done and the Values We Hold. I am, therefore, unsurprised that an educator who prefers to "grade really hard" on student papers to scare them into doing better might not see the outcomes with ungrading they desire[11]. If we want to know what the impacts of ungrading are, I suggest, looking at the educators who use it, and the contexts or institutions where they use it, is a worthwhile next step. Such a step is useful in dismantling the notion that ungrading is a singular thing, with singular, predictable results from context-to-context; and could provide an explanation for variation in ungrading's potential to benefit learning, across contexts. My attention to teachers' experiences in this project — the values they hold around writing/writing education that attract them to ungrading, and the experiences they have using it – – is in part an effort to answer these questions, constructing a model of how ungrading works in the classroom that is more dynamic and attuned to shifts among teachers and locations. Because

[11] This is not to say that educators who use ungrading never grade hard, or give the equivalent of Fs; some educators may, for instance, use "incompletes" — similar to an F, at least insofar as the student is typically required to redo the project to pass the course — to prompt students into improving their paper. Teasing out the differences here is beyond the scope of this project, though, briefly, I suggest that ungrading's emphasis on what Inoue calls "labor failure" — whether students have, so to speak, *done the things* required of them in the project — versus the focus of conventional grading practices on "quality failure" might in part account for it. A focus on labor failure is often healthier, as it directs attention away from students' failure to perform to the standard of white American English.

my own research occurs at a single location, the First-Year Composition Program at the University of Tennessee, I do not need to pay (as much) attention to what model(s) teachers are using; as I show later in this project, most educators use similar variations or recombinations of labor-based grading. This focus allows me to put more attention on questions of how educators' goals with ungrading affect their experiences using it in the writing classroom.

The second key theme I engage with in this project — arguably, to a greater extent — is the potential of ungrading to cultivate equity. As the history above demonstrates, ungrading has been, for educators, a tool for equity since some of its earliest iterations in the 1960s. In more recent scholarship, this interest blossoms into a focus on dismantling within-classroom hierarchies, especially between students and teachers. Shor notes that his contract specifically centers "power-sharing" (pg. 13) by giving students a say in what standards they are graded on and how the grading works. This work additionally hones students' ability to participate in a civic, democratic society, ensuring they have access to power not only during but after school as well. Equity and power-sharing, crucially, are framed not as in competition but coordination with the potential of ungrading to support effective pedagogy. By articulating the standards for "good writing" on any given assignment, for instance, students not only get a say in how they are graded but grow as writers, gaining crucial skills around metacognition and self-assessment (Inoue 2004). The overlap among equity- and learning-centered arguments for ungrading demonstrates the conversation is not either-or but is in fact large, messy, and frequently overlapping, aligning with teachers' multiplicitous values for ungrading.

Two points are worth noting about the interest in equity, however. First, by *equity*, many scholars simply mean flattening professional/institutional hierarchies; they ignore the ways students' identities give them more or less power within the institution. As I describe later in this project, universities are constructed in ways that make it easy for white men — and those who can perform a white man's writing style, using standardized American English — to succeed, and to keep out others, or interfere with their success. While discussion of ungrading's potential to minimize the teacher-student hierarchies is crucial, it leaves unaddressed the ways those hierarchies are (re)shaped by the identities that both students and teachers bring to the classroom.

Relatedly, despite interest in equity, much ungrading scholarship assumes the importance of asking students to align their writing with institutional norms for academic/formal conventions. Indeed, Shor argues for qualitatively grading student work, as part of his negotiated contract, as a way to ensure that students — many of whom are students of color, and/or from working class backgrounds — develop the language skills necessary to succeed professionally[12]. Elbow, more invested in students' engagement with writing and agency as writers, does not argue as strongly for the importance of cultivating students' professional writing skills, but by continuing to award A-grades only to those texts perceived as "good," according to his own qualitative judgment, perpetuates the notion that aligning with institutional norms is crucial for success in academe and

[12] To be fair, Shor also acknowledges that students are thoughtful, engaging writers with plenty of good ideas — they just need help articulating their content more effectively. This, however, does not address the ways that students own language may be most effectively in expressing their content; c.f. Duffy 2018, Baker-Bell 2020).

is a worthwhile goal to aim for[13]. In other words, texts on ungrading invested in equity challenge existing institutional structures but without critiquing the structures associated with identity and language use interwoven with those institutional structures, taking for granted the value of norms around standardized American English and assumptions about what counts as good writing and who can produce it.

Both of these points are in part addressed in the work of Asao Inoue, who argues for ungrading – – specifically, labor-based contract grading — as a form of writing assessment that supports raciolinguistically diverse students and makes it possible for them to draw on their rhetorical and linguistic traditions in composing. Inoue's work is perhaps the most dominant work in the conversation on alternative forms of writing assessment, at both the classroom and programmatic level, so far. Judging from his Google Scholar page, Inoue has more than twenty books, articles, blog posts, chapters, or other pieces on alternative forms of writing assessment, in a field where many (though certainly not all) scholars engage the subject as a secondary interest and may only have one or two. Additionally, Inoue's work — along with that of his sometimes-collaborators, among them Mya Poe and Norbert Elliot — shifts the grounds of discussion from ungrading as a way to promote learning and work towards dismantling classroom hierarchies to ungrading as key to constructing antiracist writing spaces. In this way, Inoue's work moves the conversation about the effects of ungrading in writing classrooms a step forward, onto new ground.

Conventional grading practices, as Inoue explains in his two books, *Antiracist Writing Assessment Ecologies* (2015) and *Labor-Based Contract Grading* (2018), belong to educational environments shaped by a "white racial habitus" — in other words, in the systems of values and assumption that privilege white ways of speaking and making meaning. When we grade student writing, we compare, or rank it against, standardized American English (SAE), which emerged in the early 1900s alongside eugenicist claims that literacy —specifically, *white* forms of literacy —signified intelligence (Elliot). Inoue specifically proposes labor*,* or the work that students do in the class, suggesting that labor or work is a metric more available to all students (not just white ones), as a way to ensure the "full range of grades" (3) is available to all students and cultivate value — both external, as students can use the grades for access into scholarships, graduate programs, and careers, but also internal, as students, working without the imposition of white language norms, may find things to appreciate about their own writing habits. By offering alternatives to valuing student work based on their performance of SAE, or singular notions of "academic" or "formal" English, writing educators can dismantle those value systems and invite students into a discussion of writing and meaning-making in the classroom. Inoue's work is key here, as it opens space for educators to take up Black language pedagogies (Baker-Bell, 2020) or other antiracist forms of education. Affirming students' home Englishes as linguistically valid (not "slang") and rhetorically powerful during class, then turning around and grading students using systems constructed around singular, white American English, is pedagogically incoherent. Writing educators, by offering alternatives to valuing student work based on their performance of SAE, or of singular notions of "academic" or "formal" English, can dismantle these value

[13] Interestingly, the emphasis on ensuring students have access to standardized forms of language as a way of preparing them for academic and professional success echoes concerns raised by Sherri Craig, who suggested in her Cs talk (Barrios et al 2023) that what Black students who grow up in disadvantaged communities — an identity she shares — needed was to learn to write. I address Craig's argument more fully below.

systems and make room to discuss and try out other ways of writing and meaning-making in the composition classroom.

Clearly, Inoue's scholarship is a powerful, important contribution to the conversation. But the longer I worked with it over the course of this project, the more apparent two interrelated limitations became to me, constraining writing teachers' and administrators' ability to take up and employ antiracist forms of assessment within contemporary educational systems. First is that, as with scholarship on the impact of ungrading on students' learning, Inoue similarly focuses on the impact of ungrading on the racial dynamics within the classroom — yet pays correspondingly less attention to the question of how classroom and programmatic dynamics impact ungrading, working from the components of the classroom, including the teacher and the institutional framework, in towards the ungrading practice.. This is not to say that Inoue ignores this second question altogether. Indeed, working within the framework of assessment ecologies, or (put simply) the idea that assessment is not a discrete thing that happens but is in fact part and parcel of an interconnected classroom space, Inoue is not unaware that grading not only shapes but is shaped by other elements of writing instruction, among them teachers, students, and administrators. He acknowledges that both teacherly and institutional goals matter in determining the impact of a particular assessment framework, especially one intended to be antiracist (pg. 134, 2015); and he explicitly includes "people" as a key part of the wider assessment ecologies, not only shaped by but shaping writing classrooms. Still, in discussing the role of "people" within the ecology, Inoue's focus is primarily on students, not teachers; and as his research continues, he maintains a frame on how labor-based assessment as a "grading practice builds equity and inclusion in diverse writing class-rooms [sic]" (p. 21, 2019). The role of other people — teachers or administrators — in constructing antiracist assessment practices receives less attention.

I want to be careful here. Inoue's extended attention to the transformative potential of labor-based assessment and other forms of ungrading is warranted, as an invitation to teachers to (re)consider the standards they use to judge student work and strategies to invite students' home Englishes into the classroom as rhetorically powerful. Yet, the longer we use ungrading and think about its effects in our classroom, the more we need to question how we as teachers, and the spaces where we work together, exert influence over the outcomes we see from our ungrading practice. I suggest that a more robust conceptualization of assessment ecologies can help us as researchers and teachers make this transition.

Closely related to the emphasis on ungrading as the primary vector of change is Inoue's conceptualization of assessment ecologies — a theory Inoue develops from ecocomposition (he cites Coe 1975, Cooper, 1986, Dobrin & Weisser 2002) but, surprisingly, without attending to then-emergent theories of rhetorical ecologies, especially Edbauer (2011) but also Gries (2015)[14]. As it stands, Inoue's definition of assessment ecologies leans hard on concepts of networking within and across a space, stressing classrooms as interactive communities with the various constitutive elements — among them, power, people, and component parts such as

[14] I am grateful to Seth Kahn for taking the time to discuss Asao Inoue's scholarship with me and helping me refine my concerns (which I outline below) that Inoue's theory of assessment ecologies neglects the inherent "systemness" of the university. Some of the points I make here reflect that conversation.

grading and feedback — mutually (re)shaping each other. The goal, Inoue writes, is an antiracist assessment ecology, meaning:

> A complex political system of people, environments, actions, and relations of power that produce consciously understood relationships between and among people and their environments that help students problematize their existential writing assessment situations, which in turn changes or (re)creates the ecology so that it is fairer, more livable, and sustainable for everyone (p. 82, 2015).

Working from this definition, I see *assessment ecologies*, as Inoue envisions them, as a spiderweb crisscrossing the classroom space(s). Its ends anchored in concrete, constituent elements such as students, teachers, the writing curricula or assignments, and grades, the web makes each element part of the greater whole. This in turn means that each element (for instance, the meaning of grades to students) can only be understood together with other elements; and that each element exerts influence over the others, as the strands of spider thread pull the web out and give it its shape. So far, so good. This conceptualization of assessment ecologies is valuable, insofar as it stresses the interactive and inherently political nature of assessment — the reality that we cannot simply "do" assessment tasks but that assessment *is done* to us (even if we are the teachers administering the assessment), (re)shaping and exerting power over what we value in student writing and what outcomes we see in our writing classrooms. We, students and teachers alike, exist within writing assessment.

But what is missing from this definition is a sense of the interconnectedness of the classroom ecologies with other, geographically-related ecologies — the program, department, and wider institution. Inoue is careful to situate assessment ecologies as exceeding the actual physical or temporal space of the classroom, for instance extending them to include the places students compose assignments (their dorm room or the library) and the digital spaces where students interact with each other and with the teachers (such as discussion boards on the LMS). Yet Inoue's description of the constituent elements of an assessment ecologies focuses primarily on students and teachers, bypassing the interrelated institutional spheres of action which contain and exert influence over the classroom spaces and the relationships that define it. As rhetorical ecologies, a set of theories that Inoue does not engage with in his work on assessment ecologies, explains the importance of engaging the interrelatedness of multiple ecological contexts, developing a concept of *ecologies* that is less spider web — a unit unto itself — and more a set of Russian nesting dolls, mutually containing and shaping the other.

Assessment ecologies are rhetorical ecologies

I take up this work in the following section, linking rhetorical ecologies up with Inoue's conceptualization of assessment ecologies to develop a way of framing our classroom assessment to craft a framework that, as with the Russian dolls, nests classroom activity inside programmatic or even university activities and architecture, to extend our view to include but reach beyond the classroom and situate the work of ungrading within multiple, meaningful contexts. I first define *rhetorical* ecologies, showing how this theory provides for a more expansive model of assessment ecologies, flexible enough to account for the situatedness of classroom assessment — and the students and teachers who participate in and construct it — within larger institutional

spheres. I then turn to a discussion of what rhetorical ecologies adds to Inoue's work, and the conversation about ungrading broadly.

Rhetorical ecologies emerges as a new frame or conceptualization of rhetoric in the early 2010s (Edbauer 2011). This theory counters earlier sender-receive models that stress intentionality within a given rhetorical situation. Instead, calling attention to the fluid, dynamic boundaries of human communities, especially in an increasingly networked age, rhetorical ecologies downplay rhetorical agency and purposefulness and instead stress how messages emerge from, circulate through, and both change and are changed by multiple overlapping communities, institutions, or spaces (Edbauer 2011, Gries 2015) — the way, for instance, the "Obama Hope" image, originally released as a campaign poster, was subsequently memeified online, taking on innumerable new, often ironic, meanings across the multiple contexts and spaces through which it moved (Gries 2015). Applied to institutions, ecological work stresses that meaning is not found within singular locations, texts, messages, or encounters but in the whole, at once comprised of and more than the sum of its parts (Sackey and DeVoss 2011). Rhetorical ecologies, then, as a concept stresses that meaning and meaning-making are inherently and inextricably geographically and/or institutionally situated, and that the situatedness exerts influence over the meanings which emerge from the cycling and recycling of rhetoric throughout (a) given context(s).

I turn to Edbauer's discussion of "Keep Austin Weird" (2011) as a clarifying illustration. As Edbauer describes, rhetoric often escapes the field(s) or space(s) where it originally circulated and (intentionally or not) enters new spaces, assuming new rhetorical valences or directions. Describing how the slogan "keep Austin weird" moves outside of its original context — a protest against chain bookstores moving into the city, Edbauer notes the phrase comes to circulate on radio (17), in whitepapers (17), and — ironically — as "advertisement[s]" and "corporate slogan[s]" (18). Cropping up in such divergent situations, the slogan "keep Austin weird" stresses that rhetorical activity is dynamic, living, moving smoothly across contexts, at once changing and being changed by them, often without explicit intention on the part of any one rhetor or rhetorical agency. Amplified in subsequent work, theories of rhetoric/rhetorical ecologies stress their "nonlinear [and] inconsistent" nature as they ebb and flow "across multiple networks of associations" (Gries, 7). What is key, in other words, about rhetorical ecologies is not just that they are networked systems, the (largely) self-contained spider web anchored and given shape by its nodes. Rather, they are interwoven with and dependent on other ecologies, other systems or spaces, for their structure. I want to stress here both the importance of refusing to see any one ecology, such as a classroom, as self-contained, instead attending to its *systemness* — its situatedness within multiple systems stacked on top of or within each other. I also want to stress, following Edbauer and Gries, the limits of intentionality in describing activity within those systems. Instead, I invite attention, via institutional ethnography, to the ways that the ongoing activities or architecture of the program may, purposefully or not, resist or invite certain kinds of assessment.

That theories of rhetorical ecologies, as I have outlined them here, play little role in Inoue's theorization of assessment ecologies is problematic for me for two reasons. First, this gap means that the primary ecology Inoue engages as relevant to the motivations for or outcomes of

ungrading is the classroom, framing it as a largely independent or self-contained unit (a spider web) unattached from other systems (Russian dolls). Inoue does engage, extensively, the *cultural* level, situating classroom ecologies within the wider "white racial habitus" — or the cultural milieu which situates white ways of being and making language as the norm or standard, a context which powerfully constrains and shapes how we value language (2015, pg. 47-48). Inoue also nods towards, or engages in in other texts, relevant university systems such as the writing program (see Inoue 2015, "Self-Assessment as Programmatic Center.") In his work on assessment ecologies specifically, however, his focus on classroom assessment largely bypasses intervening ecologies, more concrete, such as the first-year writing program, the department, and the university. Writing classes exist within all of these ecologies simultaneously, and the ongoing activities of these ecologies matter to what activities and relationships come to constitute the writing classroom. Put another way, including rhetorical ecologies, with their keen(er) awareness of the mutually-informative nature of multiple adjacent systems, in our understanding of assessment ecologies can help explain not only how alternative assessment powerfully reshapes the classroom but also how the turning wheels of university spaces affects the shape(s) classrooms take, in ways that expand to allow for, or constrain to restrict, more liberatory forms of assessment.

Additionally, my sense is that the absence of engagement with rhetorical theories may in part amplify the importance of a concern I have already named, Inoue's focus on assessment as a driver of change within an ecology and corresponding lack of focus to the ways other elements of the classroom shape assessment. Because the classroom is situated within wider ecologies that writing teachers participate in, among them the First-Year Composition Program and the English Department, fuller attention to the ways teachers work within these spaces to (re)shape assessment is key. This is especially true if we consider other spheres of action within the university; graduate students and non-tenured faculty, who inhabit a place of contingency and uncertainty within the department and university, are likely to affectively engage with assessment in different ways than a tenured faculty member would, in ways that, again, shape and constrain the outcomes of assessment within a classroom ecology. This line of inquiry can, importantly, help explain how educators can see such widely-varying results with ungrading and alternative assessment, why Elbow and Inoue can find alternative assessment such a powerful force for change while Inman and Powell (among others) experience much more mixed results, as their experiences and locatedness within the university is key to how they interact within the much smaller, more localized space of their classroom.

In focusing my critiques on Inoue's work, I do not see myself as singling his body of scholarship out but simply as acknowledging its outsized — and largely positive — impact on the direction and conversations about (alternative) forms of assessment in writing studies. Inoue has in fact dominated the conversation about writing assessment for the last decade, though increasingly, other scholars are rising to new prominence. What this means, however, is that the assumptions Inoue makes, the way he frames assessment ecologies and the focus on ungrading as the agent of change, influence the direction the conversation takes — what gets discussed and what does not, what changes we adopt in our classroom and how we adopt them, and what critiques are raised. Most recent literature advances modifications of Inoue's work, for instance, the role of antiracist feedback practices (Wood, 2020) or the importance of flexible, student-centered participation

policies (Gomes et al, 2020). This work is useful but does not challenge the assumptions or framings of Inoue's work, leaving open questions about the interconnectedness of the classroom with the other, multilayered systems of the university and the role of those who constitute the university — among them, teachers — in (re)shaping assessment and the ends of assessment. This is work my own project aims to take up, adopting an ecological view of assessment that intentionally accounts for *rhetorical* ecologies, as a way to attend more fully to the role of institutional structure(s) and missions in directing the purpose and value for grading — and for ungrading — within the writing program and larger university.

Rhetorical assessment ecologies — so what?

Locating classroom assessment decisions — in this case, ungrading — within multiple, interrelated university ecologies is crucial, as doing so avoids the danger of depicting ungrading as a cure-all or magic pill for injustices in university education, and instead prompts a more holistic understanding of how ungrading functions and bears fruit. Sherri Craig, currently the most incisive critical voice on ungrading, argues that especially if thoughtlessly implemented, labor-based grading places students and faculty from minoritized communities at risk (2021). Presented as antiracist, labor-based grading may make the university seem like a safer place than it actually is. Labor-based grading may also make faculty in particular vulnerable to pressures such as student evaluations or grade complaints, as Craig also discussed in her Cs 2023 talk (Barrios et al 2023). Instead of ungrading, Craig urges faculty concerned about racial equity towards what she sees as longer-term, coalitional, activist work across multiple university systems.

I am of two minds about Craig's argument, that ungrading (Craig focuses specifically on labor-based grading) interferes with effective, equitable writing education. On the one hand, I disagree with Craig about the value of labor-based grading for equity-related ends. Craig implies that labor-based grading contracts may prevent students, especially Black students, from receiving adequate support in producing the kind of writing that leads to academic and professional success (2021, Barrios et al 2023). I am not certain whether Craig's argument is that labor-based approaches provide students less feedback on the quality of their writing, or that permitting or encouraging students to use their home language in academic writing is inadvisable[15]. In either case, I would — as respectfully as possible — suggest that Craig is mistaken. To the first, labor-based approaches allow for qualitative feedback on student writing; they simply do not attach the *grade* to that feedback. Instructors I speak with over the course of this project describe providing feedback to students enrolled in their courses, though in a few cases, instructors also shared Craig's concern about underpreparing students for the standardized language norms expected in subsequent courses and professional work. Here, I would suggest that if instructors remain concerned about preparing students to produce language that meets the (singular) norms of professionalism, such feedback and coaching can be offered up within the context of an ungraded or labor-based system.

[15] Barrios et al 2023 was also a conference talk — by nature, a more fluid, open-ended genre, which resists the easy, accurate recall of a published paper that can simply be pulled up, repeatedly, for respondants to double-check. Further, I attended the talk on a whim, the last day of the Conference on College Composition and Communication (2023), so I did not take notes. I welcome Craig's, and other's, critiques of alternative assessment practices, and I have summarized them as accurately as possible here. I further welcome correction if my summary is inaccurate.

To the second, linguistics research demonstrates that teaching writing (and language broadly) in ways that affirm the value of Black language, and by extension Blackness, in professional contexts is not only pedagogically but ethically sound. As Baker-Bell (2020) points out, using standardized or white forms of English has not protected Black people from violence (pg. 30-31, 2020). Even if Black students are physically unharmed, prioritizing standardized English as the only correct way to write — pushing students to codeswitch — has an impact. Requiring codeswitching forces students to "turn on and off the parts of their language that reflect Blackness" (Baker-Bell, 2021) and can harm Black students' sense of identity and self-worth. I want to be clear that labor-based forms of grading do not require students to write in their home language. Nor will all versions of labor-based grading, or ungrading broadly, successfully invite students into a place where they can experiment with, and come to value, Black language — or for students from other backgrounds, their own English(es). That Baker-Bell traces, over the course of an entire book, the ins and outs of Black language pedagogy speaks to the robust work required of teachers looking to support students' diverse linguistic identities; the support does not automatically appear, just because a teacher takes up a version of assessment that promises to be more equitable. My point here is simply to speak to concerns that Black students, or other students from raciolinguistically minoritized backgrounds, depend on training in standardized forms of language for success.

On the other hand, however, despite my disagreement with Craig on the value of labor-based grading specifically, I share her concerns about the dangers of pushing for alternative assessment — however good or beneficial or equitable we think it may be — without attention to the possible impacts it has on the teachers, and students, who use it. Her arguments, that ungrading can increase the vulnerability of minoritized faculty and gloss over harmful inequities across the university, place a finger on the reason that emphasizing classroom ecologies without a corresponding focus on the wider rhetorical ecologies of the university is dangerous. So long as researchers and writing teachers gloss over the ways that the wider programs, departments, and university systems contain and shape the writing classroom, ungrading will not adequately grapple with the ways that writing teachers' decisions, and the impact of those decisions, are not in fact discrete, independent choices that educators can make or not make, at their leisure. Those choices emerge from, and have impacts on, educators' status within the social systems of the university. As Craig, a Black woman writing pre-tenure, reminds us, educators' positions within their departments and institutions (determined by race, gender/sexuality, socioeconomic background, and experience) play a significant role in determining how, and to what ends, they take up risky new assessment practices such as ungrading. Additionally, as Craig's urging towards interdisciplinary, cross-institutional action reminds us, the classroom is not a sealed unit; implementing ungrading as an antiracist practice and then dusting off our hands, as though we have done the thing, will do no good at all. Attending more fully to programmatic structures and the systemness of institutions in our analysis of ungrading, then, does two things: first, it helps ensure that in taking up ungrading structures, then, we do so in a way that accounts for the positionality, needs, and safety of the faculty — who may be contingent, racially minoritized, or have other vulnerabilities — teaching first-year writing. It also ensures that we see the work of researching and adopting ungrading not as a stopping but a starting point, along the way towards more learning-centered, liberatory writing assessment and educational practices.

So where do we need to go next with ungrading?

Ungrading, ultimately, proves an enduring and compelling assessment practice for writing teachers and scholars. More than forty years' worth of academic interest establishes that the practice offers up two key benefits: the potential to improve learning in writing classrooms, and the promise of equity — at first, simply flattening classroom hierarchies but subsequently working towards antiracist ends by freeing students from the strictures of white, standardized American English. Despite some disagreement about whether ungrading actually *achieves* these ends, interest in the practice remains strong, sparking current articles on subjects from applying ungrading in our feedback-giving practices to grappling with the impact of ungrading on neurodivergent students. Still, as my discussion of Inoue's and Craig's work in particular shows, much work remains to be done on ungrading. Ungrading research tends to focus on the impact of ungrading on the course, and the students taking the course; and on the classroom as a (largely) independent entity, under the control of the instructor of record; much less work grapples with questions about how classrooms are situated within other institutional ecologies, such as the writing program, department, or university, and how teachers' own locatedness within the institution, and their goals in taking up ungrading, may influence the outcomes of ungrading. Addressing questions such as these — the systemness of ungrading, and its variability based on its location and the instructor using it — can assist first-year writing programs in supporting the use of learning-centered, liberatory assessment practices among contingent faculty in a sustainable way.

I address these twin themes, institutional support for grades/grading and the necessary, corresponding supports that make possible a shift towards ungrading, in this project. A three-part longitudinal case study on writing assessment in the First-Year Composition Program at the University of Tennessee, Knoxville, this project explores 1) historical data on the emergence of grades/grading within the program during the 1920s and 30s. The project then jumps ahead a century to the emergence of ungrading (2020-2022), drawing on qualitative data collected via interviews/focus groups with contingent writing faculty in the program to explore 2) programmatic supports for alternative forms of assessment and 3) instructors' individual experiences with taking up ungrading in their own classrooms.

Put together, my key question in this project asks about institutional frameworks for writing assessment and teacherly choices about assessment within those frameworks. I ask where grading, in the UTK First-Year Composition Program, comes from historically and in light of that history, how the Composition Program can support the use of alternative, more equitable forms of assessment among contingent faculty, and how those alternatives — which I call ungrading — come to have an impact in the classroom.

I come to three key conclusions in this project: 1) that historically, institutional values and priorities drive the function and use of grades and grading, prioritizing rigor and prestige over learning; 2) that a hospitable, versus a hostile, writing program architecture is essential in inviting innovative, antiracist writing assessment practices; and 3) that within a hospitable programmatic architecture, ungrading is a fluid, not a static, practice, its affordances dependent on the values and beliefs about writing that teachers bring to the classroom; relatedly, the longer educators work with ungrading, the more likely it is to have an impact in the classroom.

I outline these chapters in over the next few pages, starting with the methods chapter and then linking the claims in the archival chapter that grades/grading in the UTK First-Year Composition Program exist to meet institutional needs with my argument that certain programmatic configurations are hostile, or hospitable, to more liberatory forms of assessment and showing how instructors' motivation(s) changes the direction ungrading takes, both within a program and in their own classes.

Chapter 1: Methods of Inquiry — Case Study and Institutional Ethnography

My methods chapter provides an overview of the two key practices I use in this project: case study and, secondarily, institutional ethnography (LaFrance), showing how they work together to illuminate the relationship between (un)grading ideologies and institutional structures or ecologies. The explicit use of case study is unique in ungrading scholarship. Most work in the field is (as in case study) locally situated, within the researchers' own classroom and programmatic contexts. Yet the work is rarely acknowledged or framed as case study, making my deliberate use unique. Additionally, my use of institutional ethnography — a method of inquiry suited to case study, given its focus on understand institutional activities through the informal, everyday operations — is to my knowledge entirely new to ungrading research. Case study, a sustained, multimethodological exploration of a singular site in response to a research question, invites attention to local contexts and the relationships among those who inhabit those contexts. This makes case study suited for assessment research, as assessment often takes vastly different shapes context-to-context (Inoue 2015). With their combined interest in locality and relationship, case study and institutional ethnography allow me to situate (un)grading choices both within classroom context(s), especially as shaped by teachers' goals, but also within wider institutional spaces, showing how assessment often reflects institutional goals and is restricted — or invited — by the programmatic architecture. I also discuss my positionality in this chapter, describing how various aspects of my identity and research experience — my whiteness, my role as Assistant Director of Composition, my responsibility for kicking off the turn towards ungrading, and my newness as a researcher — shape how I see the data in this project and the results I find.

Chapter 2: A History of Assessment in First-Year Composition at UTK, 1926-1937

I then turn to the three interrelated data chapters that constitute the heart of this work. In the first of these chapters, I draw on archival data publicly available in the University of Tennessee Special Collections to explore how local institutional values around objectivity and rigor — in conjunction with the disciplinary push for standardized American English — shaped grades and grading in First-Year Composition courses from the 1920s to 30s. As I show, grades primarily functioned to rank student writing — and students — by their performance of SAE, establishing "good writing" as a Platonic ideal unreachable by most (all?) students and burnishing the university's reputation as an elite institution. Surprisingly, we know very little about the history of grades and grading (Boyd, 1998, Schneider and Hutt, 2014). What scholarship exists typically provides a big-picture view that includes multiple universities, often focusing on Ivies or elite public universities — Harvard, Yale, the University of Michigan, or the University of California, for instance (see Berlin, 1987, Boyd, 1998, Brereton, 1996, Berlin, Connors, 1997, Ruiz, 2016). An in-depth case study at a flagship university in the Southeast — the University of Tennessee,

Knoxville — allows for a close up view of the institutional function(s) of grades/grading and their relationship with, and ability to illuminate, core values that drove writing education at the university. This research supports claims made in ungrading literature that grading is not about learning or supporting student growth but primarily about ranking students by the standard of white American English (Inoue, 2018). In doing so, the work provides a point of comparison for subsequent work, both in this project and after it, that emphasizes teachers' interest in grades/ungrading as a way to depart from the authoritarian, institutional norms.

Opening the chapter with a brief discussion of the history of grades/grading, I turn to the history of the University of Tennessee, Knoxville and its first-year composition classes during the 1920s and 30s, showing how despite claims of serving the public interest, both the University as a whole and first-year composition courses (ENGL 111, 112, and 113) specifically emphasize institutional aspirations of elitism and rigor. I then turn to the core of the chapter: aggregate grade distributions year-over-year in first-year writing, based on gradebooks kept by John C. Hodges, a respected faculty member within the department. As I show, grade distributions consistently fall along a bell curve, with Cs and Ds often the most common grade. The distributions suggest a model of grades that assumes their objectivity and authority, positioning students as perpetually falling short of standardized forms of language. Because "good" writing is increasingly tied to standardized American English, or *white* versions of English, grading practices further construct "good writing" as something achievable by white men and functions to uphold the University reputation while preventing certain groups of students — racially minoritized students especially — from accessing or thriving in higher education.

Chapter 3: Interlude

Prior to turning to the final two chapters, about the contemporary (2020-2022) turn towards ungrading in the First-Year Composition Department, I provide an interlude — a short chapter that outlines the methods for collecting and analyzing contemporary definitions and defines ungrading within the context of the UTK FYC program. This chapter answers ground-level questions about the turn to ungrading — how it came about, who participated in it, and the informal and formal departmental architecture(s) that restricted — or, increasingly — supported it. Relatedly, I also distinguish the informal activity and support instructors received in taking up ungrading from 2020-2021 with the more formal, institutional sponsorship provided from 2021-2022, via the Equitable Assessment Pilot — a department-funded trial of ungrading in first-year courses, for instructors who wanted to opt in. I also show that in the context of the UTK program, *ungrading* means largely labor-based approaches, often with collaborative and/or self-assessment, that cultivate a non-hierarchical relationship in which teachers approach students as a coach, not a judge. Key to this chapter is that ungrading comes to mean a broadly labor-based approach that, by calculating students' grade via the work (or labor) they do in the course, instead of singular, qualitative judgments of student work, undermines a hierarchical or authoritative stance towards student work and frees teachers up to pursue pedagogically-sound, liberatory assessment practices. This context provides the backdrop for the final two chapters of the dissertation.

Chapter 4: The Role of Hostile vs. Hospitable Programmatic Architecture in Supporting Ungrading

The next findings chapter, on the UTK FYC program's orientation towards (un)grading, offers up perhaps the most important intervention in this project: the distinction between hostile and hospitable programmatic architecture — or in other words, institutional structures that resist or welcome innovative, antiracist forms of assessment among contingent faculty. As I noted earlier, much of the research on ungrading focuses on the classroom level — teachers' decisions about what happens in their own classrooms — without engaging questions about the larger programmatic architecture that shapes and guides what happens in classrooms. Though a few very contemporary articles engage with the programmatic level, they do so in limited ways, for instance describing a mandated all-classes shift towards ungrading within an online program (Stuckey et al, 2020). In my project, I focus on the programmatic architecture(s) that, whether intentionally or not, suppressed or sustained the use of more liberatory forms of assessment among contingent writing faculty — by far the most common instructors in first-year composition. This focus allows me to both situate classroom activity (the focus of my final chapter) within the wider institutional ecologies that contain and shape it and to offer up a sense of the programmatic structures necessary to voluntary, continued use of equitable assessment in first-year composition.

This chapter includes two sections: one focusing on the informal programmatic support provided during AY 2020-2021 and a second on the more formal, departmentally-sponsored support, via the Equitable Assessment Pilot, from AY 2021-2022. In the first, drawing on keywords from institutional ethnography, I map out writing instructors' experiences using ungrading, when the department provided informal support via mentoring, strong models, and resources provided by colleagues, such as my own ungrading folder. I show that though even informal support and encouragement, especially from peers or supervisors, is key in supporting ungrading, preexisting features of the institutional landscape may be hostile to alternative assessment, making typical work processes such as observation difficult for instructors wanting to opt in to ungrading. I then turn to how instructors perceived programmatic support during 2021-2022, the year covering the departmentally-sponsored Equitable Assessment Pilot. Through interviews, I show that a hospitable programmatic architecture — in other words, one that welcomes risky, innovative assessment decisions among contingent faculty — provides an explicit sense of permission and safety for instructors using ungrading, along with the resources and peer and supervisory relationships to support the use of ungrading. This chapter is key to the overall argument of the project, as it directly addresses not (only) the potential outcomes or affordances of ungrading in writing classes but the necessary architectural structure to sustain those changes. Put another way, by looking out from the classroom to the programmatic and university spaces that instructors inhabit, and that affect their experiences with ungrading, the chapter provides a more holistic, ecologically or systemically expansive analysis, enabling writing teachers and administrators to make savvy decisions about how best to structure their programs to support antiracist assessment — and, crucially, the contingent instructors who take this up.

Chapter 5: The Impact of Teachers' Values on Perceived Affordances of Ungrading

I conclude with a chapter that traces contingent instructors' experiences with ungrading. Situated after a chapter showing how the UTK First-Year Composition Program provided a largely hospitable architecture for instructors looking to take up ungrading, this chapter shows how the motivations and values writing teachers bring to ungrading shapes its effects in the classroom. Also, given the close link between instructors' values and the effects of ungrading, I show that the longer contingent faculty work with ungrading, the more fruit it bears in their classroom spaces. In contrast with the pattern in ungrading scholarship, to focus on *ungrading* as the primary catalyst for change towards learning-centered, antiracist writing education, I focus primarily on writing educators, exploring the way their value(s) relate to the effects or outcomes they perceive ungrading to have. This focus allows me to develop a stronger sense of the complex, multidirectional relationship between ungrading and the educators who use it that better accounts for the variable outcomes ungrading has in the writing classroom — sometimes it works, sometimes it doesn't. Additionally, coming after two chapters that link the outcomes of ungrading to the institutional struture(s) that house and direct writing education, my focus on instructor experience, here at the end of the project, situates faculty's use of and experience with ungrading within the larger institutional ecologies. This allows me to gesture towards the ways that the values instructors bring to ungrading, and its corresponding effects, ebb and flow depending on context(s) — whether programmatic, departmental, university, and geographical or social context(s). The work invites writing program and administrators to take care in taking up ungrading, focusing less on the specific details or procedures of any given assessment strategies and, side-by-side with cultivating a hospitable programmatic architecture, attend to the interaction among instructors' experiences within the wider department and to professional development and support that focuses on teachers' values and orientation toward ungrading.

This chapter begins with a brief overview of the connection between COVID and instructors' uptake of ungrading, showing that the pandemic (and the associated political crises) are a partial but incomplete explanation for the use of ungrading; in fact, UTK writing teachers' interest in ungrading stems from deeper-seated values about the goals of writing education and their own relationships with students. Drawing on the results of open, multidimensional coding within data collected from 2020-2021, I show that instructors' core values — which range from fairness, to sensitivity towards students' humanity, to autonomy and agency in making teaching decisions — correspond with particular affordances or outcomes that instructors perceive ungrading to have. Writing teachers who care deeply about empowering students, for instance, are alive to ungrading's antiracist potential; writing teachers who care about their own professional agency see ungrading as validating their labor and making their job more satisfying. I then expand my analysis to include a longitudinal angle, using data from two educators who used ungrading in both 2020-2021 and 2021-2022 to show how instructors' experience of ungrading and the affordances they perceive in it change over time. As these educators' experiences show, repeated use of ungrading, especially within a hospitable programmatic architecture, makes it possible for educators to more fully explore and deepen the values that drive their use of ungrading, leading both to more clearly apparent results in their writing classroom and attention to a wider range of results or outcomes from ungrading, including outcomes more closely related to equity.

Conclusion

My hope in this project is twofold: 1) to invite fellow writing teachers, administrators, and scholars to see more clearly what we do when we grade student writing, or prioritize administrative structures around grading and 2) to lay the groundwork necessary, in our writing programs and in our private beliefs about writing education to make things different. The grade distributions from the early years (1920s and 30s) of the First-Year Composition Program at UTK will, I hope, provide exigence for the remainder of the project, showing how grades actually push against learning and equity in ways that justify the turn, a century later, towards alternative forms of assessment. Scholarly framings of grading as "ranking" students, whether against each other or against a single standard, are often vague even for educators who have decades of experience in the classroom. (Certainly when I first became interested in assessment literature, these claims were vague for me.) By clarifying these commonplaces, my archival research makes the dangers of grading real to readers and lends exigence to the grassroots movement towards ungrading, documented in the rest of the project.

In turn, I hope the two chapters on that grassroots movement will raise discussions not only about what the benefits of ungrading are — does it sustain learning in our classrooms? does it work for equity? — but on how we get to a point where ungrading works, both as individual educators and as writing programs and departments broadly. When I took up ungrading on my own in Summer 2019, I was backed by a decade of teaching experience, confident of my teacherly commitments and my ability to make sound curricular and assessment choices. Many teachers, however, come to ungrading wanting more direct institutional support, or still developing their teaching commitments. My project explores these questions, documenting the values underpinning teachers' use of grading and effective institutional supports in hopes that this work may enable readers to make more productive use of ungrading scholarship. By reckoning with the values underpinning effective use of ungrading, and the institutional architectures needed to support this work, my project invites readers to make productive use of ungrading scholarship in cultivating more liberatory educational spaces for students and teachers alike. To be clear, talking about the benefits of ungrading remains crucial. Even now, as I talk about ungrading with teachers I meet at run club, or at a bar, I find that its outcomes are new and surprising; teachers speak eagerly of the potential of alternative assessment to cultivate risk-taking, creative choices among students.

But as ungrading literature shows, these experiences are not universal. Some educators have less productive encounters with ungrading, while other educators would very much like to do ungrading but find themselves unsure of how to develop a workable ungrading practice, or unable to buck institutional pressures oriented around grading. Given these points of resistance, my hope in this project is to get beyond the cycle of (to be a bit reductive) yes-ungrading-has-these-benefits-no-it-doesn't. I want to invite educators to instead think more broadly, and deeply — more broadly in constructing the architectures necessary for educators to make the risky, innovative choices of prioritizing alternative assessment, and deeply in reflecting on the values and beliefs about writing and writing education that shape the outcomes of our ungrading practices. In the end, my hope is that a clearer sense of the nature of grading as ranking, and of the structures and commitments that make ungrading tick, will make it possible for readers to

cultivate alternative, more liberatory ways of being in the writing classroom, and doing writing assessment.

Chapter 1: Methods of Inquiry: Case Study and Institutional Ethnography

As I discussed in the Introduction, this project explores the relationship between institutional structures and missions and classroom grading practices, both conventional and alternative forms of assessment such as ungrading. I also discuss how individual instructors' motivations and goals affect their experiences with ungrading, illustrating how ungrading is not a static but fluid practice, dependent on the teachers who use it in and the institutional contexts where it occurs. I approach these questions as a longitudinal case study, starting with the use of grades (1920s and 30s) in the First-Year Composition Program at the University of Tennessee, Knoxville, before jumping ahead to the emergence of ungrading in the program a century later (2020-2022). and the emergence of ungrading (2020-2022) in the First-Year Composition Program at the University of Tennessee, Knoxville.

I tell this story in three key parts:

- An archival chapter (Chapter 2), using gradebooks kept by John C. Hodges in the early years (1926-1938) of the First-Year Composition Program to demonstrate how grades/grading function to uphold institutional mission and notions of rigor, and enforce a Platonic ideal of good writing
- A qualitative chapter (Chapter 4) that, drawing on interviews with GTAs and NTT faculty teaching first-year writing at UTK from 2020-2022, introduces the terms hostile/hospitable programmatic architecture, documenting what institutional features resist, or invite more innovative forms of assessment
- A second qualitative chapter (Chapter 5) focusing on instructors' experiences with ungrading. Drawing on the same set of interviews, this chapter argues that teachers' values and beliefs about writing shape the outcomes of ungrading, and that the longer instructors use ungrading, the more likely the practice is to bear fruit

I include these three chapters together, despite the significant differences in time period and methodological practices, because they reflect my own journey, and the department's journey, into ungrading and research on ungrading. I adopted a partially labor-based approach in Summer 2019 and revised in Spring 2020, at the same time as I was exploring archival records about grades and grading in the FYC program in the 1920s and 30s. Studying these records convinced me of the harm conventional grading did to student-centered teaching practices and to students' development as agentive, rhetorically capable writers. Including archival records on grades and grading at UTK establishes exigence for subsequent chapters on the emergence of ungrading, by illustrating grades' function to reinforce singular, narrow standards for correct English. In the other two chapters, I trace the ongoing activity and structures — from early informal, word-of-mouth resource-sharing among GTAs to the structured, departmentally-funded Equitable Assessment Pilot — within the First-Year Composition Program that enabled my colleagues and me to take up ungrading, and contribute to its growth. I show how teachers' own values and commitments drove the outcomes they saw with ungrading, and how those values and commitments changed and deepened over time. All together, then, these chapters effectively tell the story of a shift from conventional assessment practices towards alternatives, at first

tentatively and via a grassroots movement and then with official institutional recognition, at a public, land grant R1 university in the Southeast.

Yet connecting these three chapters, with their divergent methodological demands, presents a need for flexibility. I adopt case study as the overarching methodological frame for this project, a practice defined by its attention to a single site (or case), often over a period of time, and thus well-suited for a century-long longitudinal exploration of assessment practices in the UTK First-Year Composition Program. I also rely on institutional ethnography (LaFrance, 2019) as a secondary or complementary methodological frame, particularly relevant in my attention to programmatic support for ungrading[16]. Collectively, these methods emphasize the situatedness of knowledge and invite attention to institutional practices such as assessment as ecological, or shaped by the wider frame of the university environment. This emphasis invites me to focus on an in-depth look at the history of grades and grading and the adoption of assessment alternatives (ungrading), at a particular university and at particular times. These two methods place stress on the value of thorough description, elaborating on multiple, intertwined relationships within and across the location studied, to more fully understand the distinctive experience(s) within the subject. I discuss each of these three methodologies in the space below, focusing on what these practices mean in the context of this study, their suitability for my project, and how I deploy them in the pages that follow. I save chapter-specific methodological details — for instance, how I selected and analyzed the gradebooks I use in my archival chapter, or how I coded interview/focus group data — for the relevant chapters. Because Chapters 4 and 5 cover the same time period (2020-2022), I discuss the methods for data collection in a preceding Interlude (Chapter 3). Addressing methods chapter-by-chapter allows me to provide a more immediate, context-specific sense of what choices I am making, in a space where the reader can easily access it and rely on the information to guide their interpretation of my analysis and results.

Why focus on the University of Tennessee, Knoxville?

My selection of the University of Tennessee, Knoxville as the subject of this case study is to some degree one of convenience: I pursued my doctoral work here, and not somewhere else. What this means, however, is that I was *also* here when I attended my first Conference on College Composition and Communication (2019) and, through Inoue's keynote, learned of labor-based grading for the first time. I was here when that summer (2019) I started experimenting with ungrading, and I was here when, next year (2020), the pandemic happened and I moved from a hybridized approach into fully labor-based grading. In other words, the movement towards ungrading *happened here*, and not somewhere else; there is something to study here, because *I* was here.

[16] Early drafts of this chapter also discussed the relevance of the scholarship of teaching and learning (SOTL) to my work in this project. On the one hand, the connection between SOTL work and ungrading research is warranted, as much of the literature — including this project — falls within the scope of teaching research, as it emerges out of, and informs, particular educational contexts, usually contexts to which the researcher themselves belong (see Nickoson, 2012, Nordstrom, 2015). Whether this project counts as SOTL, however, is not the same question as whether the project makes extensive or mentionable use of SOTL work — and the answer to the second question is no. Case study and institutional ethnography are much more relevant for framing and understanding my project, and the ways I make meaning in this study, so I leave discussion of SOTL and ungrading research for another day.

Still, convenience alone does not make a viable case study. (See Gomm et al, 2019[17]). In addition to my own embeddedness in the program, the UTK First-Year Composition Program is an excellent case study, as it allows me to link several closely-related data sources that speak to each other and shed light on how assumptions about what good writing looks like come to take hold within individual writing programs and how programs and contingent faculty can challenge those assumptions, by altering their assessment choices. I dig into my data sources more fully in the remainder of this chapter, and in relevant chapters, but broadly speaking, the twin data sources I access for this project are:

1. **Access to historical data on grades/grading practices in First-Year Composition.** One key piece of case studies, given their more restricted focus on a particular location and the limits that places on generalizability, is disaggregated comparison across distinctive groups within the field of inquiry (Kennedy, 1979, see also Eisenhardt, 1989 and Gomm, 2009). At the University of Tennessee, there is not only data about instructors' turn towards ungrading, there is also data about grading practices during the program's origins, in the 1920s and 30s. These are, moreover, grading records kept by John C. Hodges, already a leader in the Department at the time of the gradebooks and later its most famous Head and the author of the *Harbrace Handbook,* a guide to correct English. The fact that archival material on grades/grading kept by John C. Hodges exists alongside data about instructors' use of ungrading enables me to tell, if not a complete story, at least a fuller one, about first the growth of conventional assessment practices and then, years later, the turn towards alternatives. Linking these two data sources has advantages for a case study, as laying historical data alongside contemporary means I am able to articulate more clearly how an institution sponsors particular assessment practices — in this case, top-down, objectivist grading practices that persisted in ranking students against the then-emerging standardized American English as the sole marker of quality.
2. **Prevalence of ungrading at the University of Tennessee, Knoxville.** Prior to my interest in labor-based grading, few instructors employed assessment alternatives, certainly not in any way consistently visible to the institution or writing program, and all programmatic documents, such as syllabus templates, assumed the use of conventional grading. Because labor-based grading was uncommon at UTK prior to this research, however, there is clearer data about why and how individuals, and the larger institution, might shift towards assessment alternatives[18]. Contemporary data about instructors' use of ungrading not only offers a key point of comparison with historical use of grades/grading in the same department; it also functions as what might be called a "typical" case (Gomm et al 2009) for the audience — a clear picture of how ungrading may, largely organically, gain traction across a writing program. Against the backdrop of

[17] Gomm et al: "As Schofield has noted (Chapter 4, pp. 77–8), far too often cases seem to be selected solely on the basis of convenience, and turn out to be atypical in important respects" (5) — in other words, defined by certain characteristic features that limit or even prohibit generalization. No matter how convenient to the researcher, a case study about the use of alternative assessment practices at (for instance) Reed College, Evergreen State College, or Antioch University may be of limited value to the wider population of writing instructors/administrators, as all of these institutions have used alternative forms of assessment for some time (see Blum 2020, pg. 4).

[18] To be clear, the shift towards labor-based assessment did not happen *because* of this research. Consent forms were not distributed until nearly the end of the semester (2020-2021) or after the semester (2021-2022) and though the preliminary research had 100% participation, the subsequent 2021-2022 study had a 59% response rate, overall.

> a century of conventional writing assessment, the emergence of labor-based grading creates an ideal case for writing instructors and administrators looking to learn more about how institutions shift towards ungrading through innovative choices among instructors as a way to meet not only social exigencies but their own values and how institutions sponsor and support these shifts.

I discuss the particular features of the English Department and the First-Year Composition Program at the University of Tennessee, Knoxville below, alongside my own positionality. Here, I turn to the particular guidelines or principles which govern my use of case study in the project. Like laying my cards on the table, these guidelines make clear how the game, or the project, overall will work, articulating expectations about how I, and the reader, make meaning from the data that follows.

How does case study as a method(ology) affect my choices in data collection and analysis?

Framing this research as a case study has bearing on the kind of data I collect and on how I report and analyze that data. Three interrelated questions, or structures, are particularly relevant to my work here:

1. Generalizability, or whether and how case studies speak to locations beyond that discussed in the case? For me, this question focuses on whether and how my project speaks to the uptake and growth of nontraditional assessment methods in first-year composition programs at other universities.
2. The locatedness of research and knowledge, or how what happens, and how we make meaning of what happens, is situated within a particular time and place (Mishler, p. 2). What institutional, geographical, social, and ideological frames shape first the growth of grading in the UTK First-Year Composition Program, and subsequently, the dismantling of grading from 2020-2022 in favor of alternatives?
3. The relationality of knowledge. Case studies emphasize the importance of relationships within a single case as key to appropriately narrow, stable claims. What relationships — among teachers and the administration, from teacher to teacher — say about grading and ungrading practices at UTK?

I explore each of these themes below, first outlining the expectations established in case study methodologies, then discussing how I fulfill these expectations in my own work, or use them to move my work forward. Crucially, in many ways each of these themes is an echo or different facet of the other, so that though the following explanations overlap some, they also provide new ways of framing or understanding the work case study does to make meaning first about a single location and then, cautiously, about a wider scope of action. I also discuss briefly how case study justifies my use of multiple methods, including both qualitative and quantitative` methods.

Generalizability

One of the thorniest problems in case study research is generalizability, or the question of how results from a study may speak, if at all, to the larger population. Sociological research typically aims at a large, diverse sample size, to ensure enough variation that results transfer out of the particular people, phenomena, and/or places under consideration to the larger population (see

Kennedy, 1979, p. 662-663).[19] Yet case studies — focused on a "single unit" (MacNealy, qtd Moriarty et al, 2019, p. 124)[20] — challenge this paradigm. Case studies inherently lack the internal variation of larger studies; their inherent historical, geographical, or social limits interferes with any attempt to draw conclusions about the population beyond the case study (see Kennedy, 1979, p. 667). The point here, however, is not that nothing can be learned from case study, that it is (merely) an exercise in exploring a singular unit, without benefit to the larger population. Rather, case study researchers insist that generalization or theorization from case studies is possible, so long as researchers are clear about what generalization encompasses and savvy in their construction and interpretation of results[21] [22].

Case study calls researchers to map out the "boundaries of the case studied [and] the data collected" (Gomm et al, 2009, p. 6[23]), tracing where the landscape of research begins and ends, so that knowledge may be more appropriately situated in its relevant context(s). Identifying the area covered by the case study makes it possible to clarify what knowledge the case study offers up and what questions it addresses — and what questions it does not. Among the characteristics which constitute these boundaries are time, or when data about the case was collected and how long the case study continued; and internal population, or the characteristics of individuals who inhabit the case and their relevance to others, both at the site (but not part of the research) and outside of the site altogether (Gomm et al, 2009, p. 6). Being specific about when research took place, who was in it, and how their experience and position(ality) may be similar to or different from other people involved is key in tracing the boundaries of the case study, enabling more cautious but valid generalizations, even from a relatively small sample size. As researchers

---

[19]The push towards widely-generalizable knowledge is especially characteristic of (older) sociological and psychological research. C.f. Mishler, who sums up this approach to critique it: "Our ideal in theoretical work is the formulation of general laws, laws that we hope are universal. The essential feature of such laws is that they be context independent, free of the specific constraints of any particular context and therefore applicable to all." (2)

[20] It's worth noting here that some case studies focus on a "single unit" that is in fact quite large. Flyvbjerg, for instance, uses the entire city of Aalborg, Denmark as a single case study (p. 228). A "single unit" does not mean a small study. Additionally, some case studies may include either multiple cases, or disaggregated examples within a single case (see Kennedy, 1979). Even in these larger units, however, the question remains, how any knowledge drawn from these units or units is valid for other units, different in distinctive ways?

[21] I like Ruddin's (2006) description of the generalizations which emerge from case study: they are "a strong form of hypothetico-deductive theorizing, not … a weak form of statistical inference" (p. 800). We are not making statistically-testable claims, though, depending on the subject of our work, that may come later, merely advancing explanations which make sense of the data in our own location — and may also help make sense of the data in other locations.

[22] I like Raul Pacheco-Vega's definition of theory, here on Twitter: "Theory establishes what we can expect, and this holds for both normative and empirics-derived theory." Pacheco-Vega goes on to suggest that theory and theorization requires an "if-then" operator (to borrow from an argument later in this project, if instructors use ungrading for multiple semesters, then their sense of its goals will expand and deepen." I would suggest 1) that one role of caution around generalization from case study, through attention to the boundaries of and relationships within the case, is to construct the conditions or qualities that constitute the "if" and 2) that caution around generalization is also warranted as not all theories from case study take an if-then format. To give a second example from this project, the theory that "the affordances instructors perceive in ungrading are linked to their deep-seated professional or pedagogical values" is descriptive, not anticipatory, and more limited in form than the normative or empirics-derived theory that characterizes harder, non-humanistic forms of social science in particular.

[23] Time and subject/participant characteristics are the features which Gomm et al (2009) note. Of course, others could be considered, such as (relevant here) institutional classification or geographical location.

analyze and develop theory from case study, they rely on connections among the various points within a case to fully explore its corners, constructing a detailed, accurate map of the landscape at hand. By comparing-and-contrasting cases, or examples within a case, for instance, Eisenhardt suggests researchers may come to see their study in new, more nuanced ways, advice which echoes the broader encouragement to attend to "relationships within and across cases" (Eisenhardt, 1989, p. 544), along with the relationship between the case and the population (Kennedy, 1979, p. 666).

Researchers may also extend readers the opportunity to partake in the theory-generating process. Sometimes referred to as "transferability"[24] (Nordstrom, 2015), this interpretive shift asks how readers may draw from the detail(s) and context(s) provided in a case study and apply it to their own location, research, or population(s). On the one hand, I see the value of shifting attention to transferability, as it pushes the researcher to provide adequate description in their work so that readers may connect the study with their own contexts and provides a crucial hedge against overstating the case (Ruddin, 2006). I also like the term "transfer" for this work, as it gestures towards the ability of case studies, not to speak to the larger population as a whole, but to provide value and information for other people, institutions, or movements with an interest in the work; and when needed, I will use "transfer" or "transferability" rather than "generalizability" for this reason. On the other, I agree with Ruddin's assessment that too strong a focus on transferability as a "fix" for the difficulties in generalizing or constructing knowledge out of a (narrow) case study shifts the burden of knowledge-making from the researchers' shoulders to the readers'. Expecting readers to do the work is not only inhospitable, a poor way of relating to somebody we hope to persuade, but potentially irresponsible. As the researcher, in particular a researcher exploring her own location(s), I am the one with the most intimate, detailed knowledge of my particular subject, within my particular place(s); and it is therefore up to me to take responsibility both for providing sufficient detail about my research to justify the conclusions I come to[25].

How does the question of generalizability affect my work here?

I make several choices in this case study to ensure the validity of my conclusions. First, following case study's emphasis on ample detail, I include as much detail about the case as possible, including the institution and the individuals who agreed to participate in this study. At

[24] In this case, transfer and transferability are unrelated to the transfer of learning, as in much work on writing education (for that work, see Fishman and Reiff, 2015, and Wardle, 2007). In methodological work, *transferability* concerns the question of how readers may make use of a case study, or another similarly site-limited or narrow study. Nordstrom uses the term on the context of writing center research, and it also shows up in case study approaches (see Ruddin 2006, Watts 2007).

[25] Crucially, the researcher taking responsibility for generalizing about their own studies, rather than leaving interpretation(s) to the reader, is that this work pushes the researcher to recall where and how their work may circulate, and to place appropriate boundaries on its interpretation. In a May 2022 Twitter thread, geneticist Jedidiah Carlson documented the "misappropriation of [genetics] research" among violent far-right wing communities and urged that scholars consider this an "actualized harm" of their work that should direct their decisions, along with the decisions of those overseeing research. While this research, on assessment practices in writing education, is less likely to draw the eye of right-wing communities or be manipulated for their service, the fact remains that more specific generalization, along with explicit articulation of how research may and may not be generalized, is an ethical move, protecting against harmful or otherwise negative circulation of one's work.

times, confidentiality requires me to provide *fewer* details, for instance using gender-neutral pseudonyms for participants regardless of gender identity. I include relevant information whenever possible, however, including participants' positions (GTA vs NTT lecturer) and relative number of years teaching. I also provide as much detail as possible about institutional and programmatic contexts. Key to this work is my choice to make the name of the university — the University of Tennessee, Knoxville — public, along with the various programs (the First-Year Composition program) and initiatives (the Equitable Assessment pilot, 2021-2022) involved in the research. This detail allows me to be clear about institutional particulars such as the program's location within the larger English department and the university and the specific steps relevant to the ongoing use of ungrading; I am able to situate my findings, especially educators' relative confidence in ungrading, as I discuss in Chapter 4, within a context that makes those findings meaningful. These details allow me to, as case study researchers advise, articulate the relationship between this case and the target population(s), other first-year writing programs interested in supporting the use of alternative or equitable assessment among instructors; my hope is that in including these details, my work has value for my readers. I also discuss the COVID-19 pandemic in my study, focusing especially on its interactions with instructors' uptake of ungrading (see Chapter 5). The pandemic is certainly a distinctive feature of this case and though it does not prevent application or transfer from one context to another, it does shape the collection and interpretation of results.

Locatedness of research and knowledge

Closely related to generalizability is the insistence in case study that knowledge and meaning are inherently contextual, situated in particular times and spaces. The approach in case study to generalizability, or the potential of research to speak outside of itself to other sites, hinges on the researcher providing thorough details about the site and/or people in the study and organizing those details into a meaningful order. To the rhetoricians and compositionists the statement that knowledge and interpretation are inherently situational may seem insufferably obvious. Of course knowledge is situational; of course we know not in universals but in particular. But the role, and importance of, location is worth restating here, as a way to articulate the distinctive features necessary to case study and justify its use in this, a study of classroom assessment practices within and across a particular institution, conducted within the field of rhetoric and composition.

By the "locatedness of knowledge," I mean that case study frames its interpretations, or the knowledge generated from the research, as meaningful only as shaped by other points of reference, such as the actions and beliefs of people or the places where the study occurred. This sets case study in contrast with positivist trends in sociological research, which insist on centering, if not a universal, at least a widely-applicable theory as the desired outcome, one that stretches across multiple contexts and explains the phenomena across many populations (Mishler, 1979). In contrast, case studies, situated in particular sites, with their boundaries clearly laid out, aim at understanding and theorizing about the location in which they occur, as I explained in the section of generalization above, mapping out the contours of the landscape in sufficient detail to invite consideration of what may lay beyond. Put another way, case study is a bottom-up approach to research and theory, grounding any ideas on a solid foundation of

particular detail, which makes possible more lofty considerations about how the case study relates to populations beyond its borders.

Knowledge is located, or situated, in case study through careful attention to the constituent parts of the case. Resisting a view from nowhere, case study takes seriously the social, cultural, geographical, and institutional fabrics of the site under consideration. People, actions, spaces/places, and time all exert influence on the phenomena studied, and are essential to our understanding of that same phenomena (Mishler, 1979, Moriarity et al, 2019). In this study, for instance, we have to consider first-year writing instructors' decision to ungrade apart from the influence of the COVID-19 pandemic, nor can we understand it apart from the particular contours of the UTK First-Year Composition Program. Both of these features contribute to the landscape(s) in which the activity of interest — in this case, the turn to ungrading among first-year writing courses — occurs. The context(s) or location of interest to case study is broad and multi-faceted, encompassing researcher, geographical and ideological space(s) and place(s), chronological time, and social networks as necessary to the work of meaning-making. The goal of case study, then, is to describe or chart out the landscape of a particular context, relying on the living web of relationships which constitute that case to supply essential information and generate theory that, ultimately, may be carefully transferred to contexts that parallel or diverge from the case at hand.

How does the locatedness of knowledge go along with my study?

The inherent locatedness of case study research has several impacts on this research, both on the data collection and also on its (re)presentation in this work. Some of these impacts I have already discussed in the generalization section above, as for instance my choice to name the university where my research was conducted. Two additional impacts are worth mentioning here. One is that case study research, given its attention to the location(s) where research occurs and the people who inhabit those locations, invites me to share my own positionality and relationship to the work, particularly related the move towards labor-based grading during AYs 2020-2021 and 2021-2022. My influence and presence within the UTK English Department at the time of this research study means that I have a unique insight on the turn towards ungrading, and so, I share my positionality towards the end of this chapter.

Second, the emphasis in case study research on the locatedness of knowledge aligns it with an ecological framing of assessment practices. An approach to assessment informed by (rhetorical) ecologies, as I explain in the introduction, attends to the ways that writing classrooms are not only themselves networked spaces but situated within, and shaped by, other institutional spaces, stacked within each other like Russian dolls — the writing program, the department, and the university. The ebb and flow of activity among these different spaces is influential to determining what choices are on the table about assessment in any given context, and the shape that assessment takes. I like the emphasis in case study on the locatedness of research because it forces attention to these institutional spaces, requiring that we consider assessment not as an independent or solitary choice, within an independent classroom, but as something that is part and parcel of the ongoing activity of the university. Accordingly, in the chapters that follow, I start by focusing on programmatic structures — the relationship between grades/grading and larger university initiatives in the 1920s and 30s, and the informal and formal support provided

for instructors using ungrading from 2020-2022 — before turning to instructors' more individual experiences with ungrading as my final chapter. My attention to the locatedness of knowledge, as a nod the rhetorical, ecological nature of classroom writing assessment, also prompts the final piece of my methodological framework: attention to relationships.

Relationships

The final piece of my use of case study is attention to relationships, a focus in line with case study's emphasis on limited generalizability and the locatedness of knowledge. By tracing out the lines that link person to person, experience to experience, phenomenon to phenomenon within an individual case, researchers create a map of the wider landscape, its contour lines clearly marked for reliable navigation and interpretation — including in ways that allow for cautious application to external situations or cases. Case study researchers tend to two particular kinds of relationships overall. The first is linking two distinct or divergent data points — for instance, as Ruddin points out (2006, pg. 802), feathers and metal, as in Galileo's early gravitational experiments. That the two substances — despite their great differences — behaved similarly affirmed the validity of Galileo's discoveries. Likewise, by drawing out parallels among distinctive examples, case study researchers can develop theories that speak to the case as a whole, and to the larger population. The second strategy, the reverse of the first, asks after the characteristics that differentiate one data point from another, or variation within a project. Attending to this variation ensures the greater generalizability or relevance of the case study (Kennedy, 1979, p. 667). Asking what makes a given data point or experience stand out from the larger case creates nuance and flexibility in the theories resulting from a case study, as well as marking out more clearly its boundaries, so that it can be transferred appropriately (or not transferred) to the larger population.

These questions about similarities and differences apply not only to the data points within a given case but also to the relationship between the case and the larger population. Researchers stress the importance of considering how the case under consideration may be "typical or atypical in relevant respects" compared to the population or society as a whole (Gomm et al 2009, p. 4). The particular *term* (typical, atypical) is less important than the actual *relationship*, whether the case shares key characteristics with the population, so that a generalization from case to population is valid; or whether the case is distinct in ways that prohibit or limit generalization. Delineating the relationship between the case and its wider contexts helps researchers know where constructed knowledge may be relevant, and where it will not be.

How do relationships go along with my study?

A focus on *relationships* as meaning-making or significant is suited to this study, grounded in a conceptualization of writing assessment as inherently ecological, shaped by networks within the classroom and (importantly) among the wider institution — program, department, and university. As Inoue notes, classroom writing assessment is a networked process shaped by the mutual interactions of students, teachers, grading or assessment, assignments, and other components of learning. Add in the surrounding context of the writing program or English Department, as here at UTK, and focusing on the relationships among these various components as key to making sense of what is happening and why — in this case, the institutional supports needed for

(un)grading and instructors' experiences with it — is clear. My project is structure by relationships along three dimensions: historical vs. contemporary; institutional vs. individual; and individual vs. individual. By contrasting historical with contemporary assessment practices, I shed light on the ideological and institutional origins of conventional grading, showing where the practice — as a way to rank students by their use of standardized American English — came from, and how institutional values influenced classroom decisions. By linking programmatic and departmental architectures with individual experiences, I show how institutional architectures prove hostile or hospitable to teachers' decision to ungrading, resisting (often unintentionally) or inviting more liberatory forms of assessment. And finally, by comparing and contrasting teachers' individual experiences, I tease out how diverse motivations and values in taking up ungrading produce divergent results in the classroom, showing how ungrading is a fluid practice which varies based on teacherly goals and experiences. Altogether, attending to relationships enables me to put together a model of ungrading as an antiracist assessment practice that is deeply influenced by the surrounding programmatic architecture and the individual teachers who use it, inviting both writing teachers and administrators interested in taking up ungrading to grapple with their goals, environments, and commitments as having a bearing on the outcome of their new assessment choices.

Multiple methods

Case study research also depends on using multiple methods to explore a given case, each a possible tool in a wider toolbelt. Various methods, especially working together, are key in documenting relationships among the constituent data within a case, and the case and the larger population. Employing multiple, complementary methods can assist researchers in pinpointing interpretations "that might otherwise be hidden" (Moriarty et al, 2019, citing Kitchenham, p. 128, see also Eisenhardt, 1989, p. 538) In other words, both quantitative and qualitative data speak to each other in ways that challenge the researcher's outlook, provide course corrections, and offer new insight that allows case study to speak beyond the boundaries of its single case to other situations and contexts, as warranted. Case study researchers do not say it outright, but I see the use of multiple methods as key to the inherent relatedness and orientation towards relationships of case studies generally. By providing multiple "persuasiv[e]" and "explanatory" (Moriarty et al, 2019, p. 131) tools for exploring the data more fully, both qualitative and quantitative work helps readers visualize and map out the connections among the data points at hand. Also, given the key role which relationships play in ensuring that any generalizations are appropriately narrow and made from one case to another, relevant case(s), the power of both quantitative and qualitative work to help researchers see what they might otherwise miss (see Moriarty et al) is essential to surfacing relationships that could otherwise be missed and ensuring not only that any generalization is appropriately limited but also that the generalizations *happen*.

I use a wide range of methods in this work. I focus on qualitative work, including archival research, interviews, and some document analysis; but I also include quantitative elements. These appear in my archival chapter, where a quantitative breakdown of grade distributions year-over-year in John C. Hodges's gradebooks is key to visualizing the realities of grading during the 1920s and 30s in the First-Year Composition Program. At the same time, qualitative analysis of individuals' experiences with taking up labor-based grading in their first-year writing classes provides an in-depth, nuanced picture of instructors' use of ungrading and allows for the kind of

cross-data analysis that Kennedy and other case study scholars advocate. I work to bring out key differences and similarities from one instructor to another, not only in clearly observable details such as position (GTA vs NTT instructor) but also in less tangible, value-laden characteristics, such as instructors' teaching philosophy or their emotional response to the labor of teaching and grading. By using multiple methods, then, I am able not only to frame and explore each dataset with relevant tools but also to explore the detailed ins and outs of individual experiences, so I can more accurately connect these experiences to the wider population. I discuss each methodological choice in the chapter where it is relevant.

Use of institutional ethnography

I also draw on institutional ethnography (IE) in shaping this work. A method particularly suited to understanding actions in institutional context (LaFrance, 2019, p. 23), IE assists in thinking productively about the relationship between individual instructors who took up ungrading and the larger institution(s) of the writing program, the English Department, and the University of Tennessee, Knoxville. The emphasis institutional ethnography places on the connection between "individual experiences" and "institutional context" (Wootton 2022, p. 60) makes it valuable to this work, particularly since instructor experiences during AY 2021-2022 is characterized by explicit institutional support for labor-based grading. A methodology which invites me to attend to the way that an institutional framework at once is constituted by and (re)shapes the activity of people who inhabit it, depending on their positionality and power within that institution, is key. Institutional ethnography asserts that:

> Individual experience, ideals of practice, local materialities, and institutional discourse are mutually constitutive. . . . Using IE to study the "work" that people carry out allows writing studies researchers to reveal the deep and often hidden investments and experiences of those people, making visible the values, practices, beliefs, and belongings that circulate below more visible or dominant discourses (La France, 2019, p. 5).

In other words, I see IE as reaffirming that human choices are made, not in a vacuum, nor even on the basis of individual "values, practices, [and] beliefs," but within particular institutions, which impose their own beliefs and ways of being on the people who inhabit that institution. Yet at the same time, IE gestures towards the power of human choices to (re)shape institutional activities and commitments. Overall the emphasis within IE is on the ways that individual choices are shaped by "active social and professional norms" — in other words, *institutional* norms, including programmatic norms — "in local settings" (LaFrance, p. 25). Yet the fact that IE *also* frames institutional expectations or ways of being as negotiated and adapted by individuals who dwell within it suggest that the institution may, in fact, be changed *by* those individuals (see pg. 29)[26]. This in turn calls attention to the way that institutional activity and individual activity interrelate and shape each other. I take it up here to be clear that instructors' decisions about ungrading are both inevitably shaped by the larger institutional moves but also that instructors' choices have the capacity to reshape the institution. Reflecting the framing of

[26] While I cite LaFrance here, in fact an initial draft referenced a tweet, since lost to time. The important point, however, is that I am not the only or the first one to look at IE and see a method that effectively frames the impact of individual choices on the institution, rather than the other way around.

assessment as ecological, the use of institutional ethnography as a (partial) frame for some sections of this dissertation provides crucial terms for seeing and articulating the importance of institutional belonging in supporting instructors' experiences with alternative assessment. In this way IE also speaks to my goal in this dissertation to explore how institutional structure fosters (or does not foster) instructor agency and choice related to their assessment design, and, simultaneously, how instructor choice may in fact reshape the institution.

I specifically draw on IE in Chapter 4, on programmatic experience, or the informal and formal supports provided by the Composition Office for instructors who chose to use ungrading. There, I borrow several key terms ("boss texts" and "work/work processes" in particular, see LaFrance, 2019) and emphases (the value of relationships within and across the institution) for analysis. I reserve definition of terms *for* those chapters, as they become relevant. In this chapter, I want to stress that though IE is not my primary methodological lens, it is in fact important, drawing attention to institutional and individual practices as interrelated, mutually reinforcing activity systems. Particularly as part of a larger case study, an IE framing directs descriptive, thorough attention to the interactions among individuals and the larger institution(s) they inhabit, laying the groundwork to offer conclusions about how particular relationships or distinctive experiences at UTK affected the use of ungrading in FYC at UTK and, potentially, to put forward valid generalizations as well[27].

Limitations to this case study

I see the limitations to this case study as belonging to two categories: institutional and individual; both the location (the First-Year Composition Program at the University of Tennessee) and my own identity and status as a researcher constrain the project in ways that affect the knowledge constructed from it, and the relevance of that knowledge to other contexts. I discuss both categories here, describing the institutional and social contexts before turning to more information about who I am and how that affects my approach to this project.

Institutional Limitations: Distinctive Features of the UTK English Department, 2020-2022

As a case study, this project is inextricably situated within the First-Year Composition Program at the University of Tennessee, Knoxville. The FYC Program is part of the larger English Department, which has three primary divisions: Literature, Criticism, and Textual Studies (the largest); Creative Writing; and Rhetoric, Writing and Linguistics (RWL). Typically a faculty member in Rhetoric, Writing, and Linguistics serves as WPA. Until AY 2020-2021, the WPA

[27] I do want to be clear that valuable as IE is to my methods, I do not identify this work as an institutional ethnography, in part because it is not an ethnography. Certainly the archival chapter is not ethnographic at al. Also, I pass over some sources both terms, such as ruling relations, and sources of data, such as texts, key to institutional ethnography. Ethnographers such as Wootton (2022) insist (in a tweet) that "if you're not looking at texts, you're not doing IE". Had I conceived of this work as institutional ethnography prior to beginning analysis on the first set of contemporary data (AY 2020-2021, analyzed Jan. 2022), the key terms from this methodology may have served as a valuable heuristic for analysis. By the time I realized this, however, in August 2022, analysis was far enough along that it was not worth the trouble. I hope to explore these intersections in further analysis of the data. Subsequent research would benefit from addressing the institutional ethnography framing more thoroughly, exploring more directly how ruling relations around grades and grading, for instance, are negotiated by instructors as they take up, develop, adapt, and continue to use labor-based assessment or other grading alternatives.

role was filled by a single RWL faculty member, indefinitely; beginning in AY 2021-2022, the WPA role was converted into a three-year term, shared among several tenured faculty members in RWL. First-Year Composition is a two-semester sequence, with the first class (ENGL 101) focusing on rhetorical analysis and persuasion of diverse audiences and the second (ENGL 102) on research writing, including qualitative, archival, and secondary forms of research. Both courses typically include multimodal writing. The curriculum has been in place since Spring 2007 (Fishmann and Reiff, 2015). The UTK English Department and First-Year Composition Program has autonomy over curricular and assessment choices. Instructors are free to develop their own assessment practice, so long as it is fair[28]; any more substantive changes pass through the Composition Committee, composed of the Composition Office (Director, Associate Director, and Assistant Director), the Writing Center Director, and elected part- and full-time NTT faculty and an elected GTA. To my knowledge, there is no state-level oversight of the curriculum.

The As of AY 2022-2023, the First-Year Composition Program consists of approximately 90 instructors — a mix of graduate teaching associates (GTAs) and non-tenure track (NTT) faculty teaching approximately 200 sections of the courses. Importantly, because First-Year Composition belongs to the English Department broadly, it has access to English Department funding, in particular the multimillion dollar Hodges Fund, an endowment started with funds from English professor John C. Hodges' publication of the style manual, the *Harbrace Handbook*. The Hodges fund supported the Equitable Assessment Pilot in AY 2021-2022.

Key to the institutional context is the wider social context, often the origin of changes within university ecologies. Two key, interrelated changes are important to bring up here: the COVID-19 pandemic and the resurgence of the Black Lives Matter movement. COVID-19, or COVID, as I refer to it throughout this project, is a novel coronavirus with respiratory and cardiovascular symptomps that, first documented late in 2019, was announced by the WHO as a pandemic in mid-March 2020. In the United States, this announcement prompted widespread closure of in-person services and service work — a shift which led both to substantial job loss within service sectors, and to increased health risks for people working in so-called "essential" industries. On college and university campuses, teaching and learning suddenly shifted online, relying on LMS platforms such as Canvas and video conferencing software such as Zoom to connect students with teachers; the sudden shift strained teachers' resources and creativity and students, often forced out of dorms or apartmets and back home, often found themselves without the resources necessary to thrive at school. Many teachers, as I discuss later in this project (Chapters 3, 5) felt a strong sense of responsibility to adopt more flexible, supportive pedagogical practices to help students succeed during this time. Universities gradually moved back to in-person classes over the course of the next few years; at the University of Tennessee, Knoxville, many classes (though not all) had a significant in-person component by Fall 2021, a year after the initial closure, and by Fall 2022, most classes at UTK were entirely in-person.

[28] When I first adopted a version of labor-based grading in Summer 2019, I cleared my choice with the then-WPA, who told me via email that I could use my discretion in the assessment model I used, so long as my choice was fair and transparent for students. I base my assertation that instructors have a great deal of autonomy in their assessment choices in part on this experience.

The COVID pandemic was only one of two crises in 2020, however. That spring, the murder of Breonna Taylor in March 2020 and George Floyd in June 2020 prompted the resurgence of Black Lives Matter. Nationwide protests for racial justice and against police brutality sprang up that summer, capturing widespread attention and support, perhaps especially since many more people, out of work or working from home because of the pandemic, had the time and opportunity to pay attention. Additionally, given that many service workers and workers in essential industries, such as child care or the trades, are from minoritized backgrounds, the financial and health risks of the pandemic served to amplify longstanding racial inequities within the United States. Universities responded to the Black Lives Matter protests with programmatic or institutional initiatives designed to identify ongoing injustices on campus and to invite faculty, staff, and students to reflect on their own racial identities and pinpoint areas for change. Here at the University of Tennessee, Knoxville, the push for greater attention to racial justice in academic contexts took the form of an Academics 4 Black Lives training and a departmentally-sponsored series of talks, including my own, early research on secondary literature around ungrading, on integrating antiracist choices into our scholarly work.

Crucially, these twin crises exerted an impact on the shape and goals of the institutional contexts, pushing me and my colleagues to acknowledge the harms, especially along racial and socioeconomic lines, that grades/grading did to students and to take up alternatives that promised to be at least a step more equitable. The crises, in other words, opened space for widespread institutional change, at least along a grassroots level, that would likely not have happened, or at least would have happened much more slowly, in ordinary times. The flip side of this reality, however, is that as institutional and social practices have returned to normal in the following years, appetite for change has dwindled. The bureaucratic gears once again grind slowly. Writing instructors and administrators who come to this project hoping to learn how to support alternative forms of assessment in their own contexts will find, I hope, a particularly clear and powerful example of institutional change — but their own work to support similar change, in their own contexts, may be limited by the lack of social exigencies or emergencies to prompt that change. The substantive turn from a century-old practice of grades/grading towards ungrading that I document in this project is likely the product of a unique historical moment and cannot be duplicated or even approximated, though I hope the turn may still provide readers good food for thought.

Individual Limitations

Also important to address are individual limitations, the features of my own positionality — relative to this research project, and to the wider social contexts I inhabit — which shape the work. One key limitation is the inherently solitary nature of this project, as a dissertation emerging from my own teaching and service work. While case studies often lean on additional cases to reach more generalizable findings (see Moriarty et al, 2019, p. 131), this was not possible for this work — a dissertation, and so by nature, an independent project, which I completed largely alone. This is especially true at points where I may have benefited from collaborators, such as initial coding. Further work with a team of researchers may cast new light on my findings, or interpret them in new ways.

The solo nature of this work also extends to my relationship with participations. As Moriarty et al (2019) describe, case studies can and should involve participants beyond a subject role, attending to and valuing their input in ways that allow them "to become spokespersons for your research findings over time" (131). Without getting too into the weeds of particular practices, if I were to design this work from scratch, I would have written in a formal return to, or member checking with, participants as I analyzed and prepared their data. I did not, and though retroactively changing a study design *can* work well, as I describe in the paragraph below, it can also be chaotic and burdensome for participants and researcher alike. As it stands, then, my work is more one-sided than I would like; it represents *my* interpretation of the origins of grades and grading at UTK and, subsequently, the grassroots movement towards labor-based or other equitable assessment alternatives. My participants' interpretation(s) may differ (though I have done much in my power to ensure accuracy, as I describe in this chapter and elsewhere) but this is *my* interpretation.

Finally, the work is missing data that would make it more complete and robust, including (though not limited to) survey data from students enrolled in courses in which labor-based or another equitable assessment practice was used. While I intended to collect this data, life interfered, and I did not. Nonetheless, as Eisenstadt (1989) writes of case studies, "it [is] legitimate to alter and even add" — and here I would suggest, by extension, *subtract* — "data collection methods during a study," as "investigators are trying to understand each case individually and in as much depth as is feasible" (539). On the one hand, talking with students or surveying students about the use of labor-based grading is absolutely essential; we cannot move forward without a fuller understanding of how such assessment practices are received by real students, in our real classrooms. But on the other, by constricting the research, I landed on a project better geared to the location I was studying, the time I was doing that study — during the waning months of the pandemic, and to my own status as a beginning researcher.

Positionality

In their work *Race and Research Methods*, Lockett et al (2021) address the ethical importance of positionality, urging researchers to remember that there *are* no nonracial spaces. Pretending otherwise ignores that whiteness too is a racial category. If white researchers assign it a neutral or natural status in contrast with the racialized status of Black or Brown researchers, they replicate discrimination within academic research and practice. Researchers are instead called to directly acknowledge the way race and racialized experience(s) shape the landscape of the research and its conclusions (Lockett et al). If we do not account for the impact of race and racism on a given case study, or our role in collecting and analyzing data from it, we risk a distorted picture of that study, one that bypasses the lived experiences of participants and the currents which — whether those participants are, like me, white, or whether they are Black or Brown — shape every feature of those lives, including the features affected by the case study. Especially given the commitment of case study research, along with practitioner inquiry, to the locatedness of knowledge, I would extend this ethical commitment to call researchers, like me, to a frank discussion of how our identities, beliefs, and experiences interact with the work that we conduct and shape the form that it takes. For these reasons, then, I take time here to outline how my own identity and space in the world affects (affected) this project.

I want to focus on three aspects of my identity:

- My racial identity. I am a white woman
- My standing as the Assistant Director of Composition
- My identity as the originator of ungrading in this form

Each aspect listed here affects everything about my project — what data was collected, how it was collected, what analysis was done, what conclusions I came to and what conclusions I didn't. Mapping out the specific impacts is, of course, impossible. If, for instance, my role in kickstarting the 2020-2022 turn towards ungrading at UTK masked or hid certain conclusions from me, then I cannot (easily) say what they are; the whole point is that they are hidden from me. Instead, then, in this section I briefly trace what my identity is relative to the project and how it influenced my approach to the research, as a way to locate myself within the landscape of my study and enable readers who occupy other places in life to connect what they read to their own contexts. Transfer of this research to other context(s), such as other institutions, demands a thorough understanding of the contexts the researcher inhabits. I meet, in part, this need for understanding through my positionality statement below.

I am a white woman.

In my introduction, I noted that the racial composition of UTK faculty and (graduate) students and its status as a predominately white university affects how I situate this project within the conversation on antiracist assessment — hopefully, but with a keen awareness of our shortcomings. Here, I want to briefly discuss how my own racial identity affects the shape this project takes.

I am a white woman, raised in evangelical Christian spaces where ignoring race, often called colorblindness (rather than antiracism) was considered the best approach to countering prejudice[29]. Now, as someone committed both to honestly acknowledging my own privilege and working against injustice, I have mixed feelings about how my racial identity intersects with my research into ungrading, particularly given ungrading's reputation as an assessment practice suited for addressing racism in assessment (Inoue, 2019). As I took up ungrading within the FYC program, shared it informally with friends and colleagues, then coordinated a series of workshops sponsored by the English Department, I became the departmental expert on antiracist assessment practices — an ethically-fraught place to be. Over the last thirty months, since January 2021, I have learned and grown in my own critical engagement of students' linguistic and rhetorical diversity in my class. Currently in the general education composition courses I teach, I directly address the intersections of race, identity, and our conceptualizations of professionalism and professional writing, assigning Hull et al (2019), along with public resources on linguistic diversity, to spark students' thought. For a variety of reasons, I do not focus on these questions as much as I would like to or think necessary.

[29] Despite the emergence of color-evasive as a potential anti-ableist term for this concept, I use colorblindness as the more direct, and therefore more effective, term. (For the use of color-evasive, see Angela Davis: https://youtu.be/6D4aZoqDqOQ; for color-blindness, see Bonilla-Silva 2009. I am also indebted to my advisor Dr. Jessi Grieser for suggesting this line of thought.)

I also acknowledge the ethical complications of undertaking antiracist work as a white woman. Too often, white scholars who research issues of race and racism, and people who belong to racially minoritized populations, wind up reproducing colonialism, by claiming issues that they have no stake in for their own and using their work to amplify their social status (Lockett et al, 2021). White scholars also run the risk of forwarding solutions that do not solve the problems they intend to solve. Sherri Craig (2021) warns against well-intentioned white administrative leaders doing harm by imposing assessment methods assumed to be antiracist (in this case, contract grading) on faculty, especially faculty of color. Likewise, James Eubanks (C&W 2022) points to the dangers of administrative calls for equity without listening to racially minoritized members of the department or making decisions about departmental structures and initiatives without soliciting input from those same members of the department. These concerns are serious, and I am haunted by them, especially since my work on antiracist assessment practices has had tangible professional benefits for me, including not only this dissertation but also a grant and ongoing professional development as an administrator and expertise in assessment. I believe that an ongoing commitment to reflexive interrogation of my own practices, to ensure that I am not acting in an ignorant, discriminatory or (inadvertently) harmful way is critical.

Yet on the other hand, my whiteness creates an obligation to actively work towards antiracist ends. In Summer 2020, amidst first the pandemic and then the resurgence of the Black Lives Matter movement, I increasingly saw labor-based assessment as an urgent — if partial — response to the exigencies of social and racial injustice in the classroom. Using labor-based grading did not solve these problems automatically, but it did push me to think and work explicitly to make my classroom supportive for all students. Since then, as I have continued to work and talk on labor-based grading, I have threaded the needle carefully, at once acknowledging that ungrading is not automatically antiracist and using assessment as an opening to more directly discuss questions of race, racism, and (linguistic) discrimination among colleagues. Though, indeed, I am a white scholar participating in equity work in my department, the reality is that we are a small program — as I write this, in early Summer 2022, we do not even have ten doctoral students in the entire rhetoric and composition division — and there are no doctoral students of color in our English Department doing assessment research to whom I could defer[30]. Moreover, in our discipline more broadly, many white WPAs do not see a need to address the intersection between race and racism and assessment, or white WPAs require education about race and racial issues, a demand which exacerbates racism within institutions (Garcia de Mueller and Ruiz, 2017). For me, then, my expertise and interest in ungrading creates an obligation to advocate for equity in assessment practices.

In fact, much antiracist and equity work in education depends, for better or worse, on white women. As Moore et al (2017) point out, the reality that white women are by far the most common demographic in education sharpens their obligation towards antiracist work: “White women are in the driver’s seat in the classroom in America. We don’t have an option here. White women have to do this work” (p. 53). Put another way, so long as white women fill teaching and leadership roles, white women are responsible to engage in and move antiracist work forward;

[30] Our department has struggled to recruit and retain faculty and graduate students of color. As of this writing (Spring 2023), the department has hired a scholar (Jamal-Jared Alexander) who researches antiracist admissions practices, however, so perhaps the demographic makeup of our department will change in the future.

opting out of antiracist work only means that the labor goes undone, or falls heavily on the shoulders of BIPOC educators and scholars. As I pointed out in the Introduction, first-year composition educators at UTK, my colleagues, are predominately white; of the past five Assistant Directors of Composition (a graduate student position), all have been white women. My own interest in and expertise with equitable or antiracist forms of assessment, then, means I am well-positioned to take the lead on advocating for more widespread use and support of these forms of assessment, despite the fact that I am not myself a member of a minoritized racial or ethnic group. As the main originator of ungrading, especially in first-year composition courses, at our institution, I am "in the driver's seat" — positioned to, and therefore responsible to, work for change in writing assessment practices. The work is far from complete, of course. Much work remains to be done on supporting educators' critical, direct engagement, in writing assessment, with the use of diverse linguistic and rhetorical practices, and on constructing culturally sustaining composition curricula (Paris, 2012). But my hope is that in expanding the use of ungrading at UTK, we take a (baby) step towards more antiracist practices. I hope in this project, a study following up on that expansion, to take advantage of my identity — as a white woman in a position of leadership among first-year composition educators — to support this work, advocating for, and researching, assessment practices that blunt the impact of bias and make more room for students' own language and rhetoric.

To me, then, and in consultation with my advisors, who would absolutely call me out for discriminatory or harmful behavior, it seems more important to move forward cautiously than to stop work altogether. That said, I take several steps in this project to ensure that my work is as equitable and accurate as possible. The first is of course my close work with several BIPOC faculty on this project, notably Drs. Lisa King and Jessi Grieser, who provided guidance on framing, research, and advocacy that accurately represents ongoing equity work in our division and helps, rather than harms, BIPOC students in composition courses. Additionally, I seek out scholarship published by BIPOC scholars, not only Asao Inoue but also April Baker-Bell, Sherri Craig, Staci Perryman-Clark, Genevieve García de Müeller, Iris Ruiz, and others. These works sharpened my understanding both of the harms that conventional forms of assessment pose to raciolinguistically minoritized students, boxed into ways of being and writing, or judged against standards; that are not their own — while also highlighting potential dangers in alternative forms of assessment, among them the potential to ignore crucial cross-institutional collaboration on equity issues (Barrios et al 2023) or the greater demands that the practice places on the relationship between faculty and students (McCloud 2023). My goal was to ensure that I heard both from supportive and critical voices on the antiracist aspects of equitable assessment. By engaging with and valuing perspectives other than my own, inextricably shaped by my own whiteness, I hope to ensure more accurate, careful work, transferable to a wider range of context.

I am the (an) originator of ungrading / labor-based grading in this form at UTK.

When I first started ungrading, in Summer 2019, I was the only person I knew in the program using this model to assess students' work over the entire term. (My impression, from conversations I had around that time with the WPA, is that labor-based approaches may have been used outside of first-year courses, or on the level of a single assignment, but not for the whole course.) Now, as I write this in Spring 2023, the landscape has changed. This research

represents 11 instructors who used ungrading for their whole course, not just one assignment, though as many UTK educators who use labor-based grading are not included in this project, the number is undoubtedly higher. Anecdotal reports indicate that the use of labor-based grading has spread beyond First-Year Composition, as other members of the English Department or instructors at the University of Tennessee Knoxville share with me that they or a colleague are using ungrading.

I feel responsible for this shift. I created and freely shared resources documenting my approach to labor-based grading in ENGL 101 and 102, along with additional readings and suggestions for designing an ungraded Canvas course; I shared, at Dr. Katy Chiles's invitation, about ungrading with the English Department in November 2020, I shared about ungrading in pedagogy courses (ENGL 505, taught during 2021-2022 by Dr. Lisa King) and the start-of-term August Workshops, and with my advisor and the Director of Composition, I planned, applied for, and successfully helped administer an ongoing pilot of equitable and alternative assessment practices in first-year composition. I want to be careful about overstating my role. Other educators proved equally or more effective evangelists — Jamie, for instance, who describes campaigning for ungrading among other graduate students. But there is no doubt that I stand near the center of this change.

I am proud of the changes that I, and my fellow graduate students and lecturers, have led. As subsequent chapters describe, ungrading or labor-based grading is beneficial for instructors, making it possible for them to pursue equitable pedagogical practices more effectively and to spend work time on tasks that feel meaningful. My involvement in the emergence of ungrading at UTK, however, affects how I analyze and discuss the data in this project. Because I shared my ungrading model widely, and because (as I discuss in Chapter 4, later), it was taken up widely, the choices participants in this study make about their assessment models — choices which I discuss and which have outcomes I analyze — reflect, sometimes quite closely, my own. In analyzing participants' choices, I try to strike a careful middle ground. On the one hand, though I may discuss the outcome of one assessment choice or another, I avoid singling instructors out for blame or criticism, as my own choices, and my ability to effectively guide fellow instructors in using ungrading, are implicated. To be clear, instructors made their own choices about ungrading; I often did not know exactly what those choices were (or at least I did not pay much attention to them) until research began; and sometimes, I disagreed with their choices. At the same time, however, instructors had access to more or less information and coaching about ungrading, depending on where I was in my own ungrading journey, and I acknowledge that my presence in the Department and conversations with people about ungrading inevitably shaped how they made decisions about their assessment practice.

I do not believe that absenting ourselves from the communities we work in and study is required for ethical research. What I do believe is that part of what it means to be involved with our research, and to be an ethical teacher, administrator, and researcher is to acknowledge our locatedness within that community, our influence over the direction(s) it takes, and our obligation and responsibility to the people we work alongside and influence in heading up a program. The question is not how this person or that experienced the adoption of ungrading; but more precisely, how they experienced the adoption of ungrading within an ecology that I had a

hand in creating? Determining the exact boundaries of my influence is, of course, impossible, though I aim to honor my own presence and influence within the UTK first-year composition community by documenting, whenever possible, the steps I took to advocate for ungrading and to (re)structure the program in ways that made the choice to ungrade easier for instructors. As I note in Chapter 4, for instance, I rewrote and distributed guidelines to ensure that teaching observations validated and provided relevant feedback on ungraded courses, acting on information from my initial data (2020-2021) — a choice which undoubtedly contributed to the fact that in 2021-2022, we did not, to my knowledge, have teaching observations that criticized instructors for the choice to ungrade. At the same time, I honor educators' agency and choice in ungrading, and I center their experience in the chapters that follow. This is not, ultimately, my story but theirs.

I served as Assistant Director of Composition.

I served as the Assistant Director of Composition (AD) during AY 2021-2022, which included the time period when English, through the Hodges Fund, supported labor-based grading and equitable assessment among teachers of first-year composition. I describe the Equitable Assessment Pilot in more detail in the Interlude. Here, I want to acknowledge that my position as AD at once gave me the ability to ensure institutional support for ungrading but also gave me a greater stake in the project than an outside observer.

The AD role is conventionally held by a graduate student in their fourth or fifth year. The role centers support and coaching of instructors and is not supervisory. My primary responsibility was heading up the design and delivery of teaching workshops prior to the fall (ENGL 101) and spring (ENGL 102) terms: I created presentations, invited fellow instructors to share about ethical teaching and effective classroom management, designed and distributed Canvas templates, and organized a virtual field trip to Special Collections. During the term, I conducted syllabus review, hosted office hours via Zoom, and fielded questions about students and institutional expectations from instructors. I also conducted three observations of fellow teachers, though none of the teachers involved were eligible for the research represented in this dissertation; at no point did I hold a supervisory position over the participants in my study. Also, though other Assistant Directors have conducted IRB-approved research while in this position (Mobley Finn 2018), I took several additional steps to ensure that research participants were free in making their choice, such as waiting until individuals receiving support for their use of labor-based grading (in the form of a stipend, via a Hodges Fund grant) and until the semester was concluded to recruit.

Serving as AD during the same year we conducted the Hodges-funded Equitable Assessment Pilot undeniably allowed me to support the Composition Office in providing robust support for innovative assessment practices. Scheduling workshops to assist educators taking up ungrading was easy and straightforward. Additionally, I was in a position to (re)design institutional processes, such as teaching observations, to better validate instructors' choice to use ungrading, work that emerged from my preliminary research (2020-2021). Additionally, the close interconnection between my own work in the Composition Program, and my research among instructors in the Composition Program, falls within the boundaries of case study, accepting

locatedness within a particular time and place, rather than a presumed neutrality, as necessary for meaning-making.

At the same time, however, especially coupled with my experience advocating for ungrading, serving as AD made me more entwined with my research and committed to its success. My role as AD means that I am, in part, responsible for some of what I found or did not find — for instance, the lack of complaints in 2021-2022 about observers who are critical or skeptical of ungrading as a valid method of assessment. Ungrading was (and is) something I had a professional stake in, which may in part account for my sense that it had positive results, and that the department, especially during the Pilot, supported it. My role as AD also shaped how I collected the research, generating anxiety about observing appropriate boundaries. This anxiety at times restricted or limited the amount of data I was able to collect and when I collected it.

Conclusion

Since this is a multi-methods dissertation, specific methodological processes vary widely from project to project, and chapter to chapter. Collecting and analyzing archival data follows very different processes than analyzing qualitative data. This chapter knits together these disparate methodological choices into the larger framework of case study — with institutional ethnography as a secondary framework — and the implications for the project as a whole. Specifically, I explain my reasons for choosing the UTK First-Year Composition Program as a case study, including 1) the ability to link archival data with contemporary on the use of grades and grading and 2) the clear emergence of ungrading from 2020-2021, allowing for a clear snapshot of what happens when a program takes up alternative assessment practices in response to instructor advocacy and social exigencies. I also address the rules, or guidelines, that govern case study — generalizability, relationships, and the locatedness of knowledge, and I discuss the impact of each of these guidelines on my work, prompting me to situate my work explicitly at the University of Tennessee, Knoxville and attend to the relationships among teachers and the program as generative or insightful. I acknowledge my use of multiple methods, particularly archival and qualitative, though I also use limited quantitative analysis across each chapter, to easily capture grades and grading data and key themes in instructors' conversations about their use of ungrading; and I gesture towards my use of institutional ethnography, which I discuss more fully in the chapter on institutional experience with ungrading. Finally, I discuss my own relationship or position relative to my project at length, exploring how my racialization as white, my position as Assistant Director of Composition (2021-2022), and finally my advocacy for ungrading shape my stake in this project and my interpretation of results.

I see methods and methodology as laying my cards on the table, or (to use a different metaphor) setting up the fencing or boundaries that define the landscape of the project. Having those boundaries made clear — here is what falls within the limits of this project, here is what falls outside of it; here is how I arrive at those arguments — governs and makes possible the interpretative decisions in the chapters that follow, ensuring their consistency and validity. I turn to those chapters now, beginning with the archival section and continuing on to two contemporary chapters, the first on institutional architecture as hostile, or hospitable, to ungrading and the second on instructors' experience. I address relevant methodological details chapter-by-chapter, using an "interlude" chapter prior to the two contemporary chapters to

discuss how I collected the data used in those chapters and establishing a definition for "ungrading." In these, I argue that at the University of Tennessee, Knoxville (un)grading practices are deeply shaped both by programmatic structure and ecologies and by teachers' own commitments. This work invites readers at other locations and contexts to consider how the shape of their writing program or department, and their own values and beliefs as writing educators, influence the use of ungrading and its outcomes in their own spaces.

Chapter 2: A History of Assessment in First-Year Composition at UTK, 1926-1938

Key to my argument in this project is that institutional architecture — meaning the framework or structure provided by university organizations, such as the writing program — plays an important role in determining the kinds of writing assessment used, and their impact on students, teachers, and in classrooms. In later chapters, I show how the contemporary First-Year Composition Program at the University of Tennessee (2020-2022) invited teachers' use of liberatory forms of writing assessment. Here, I lay the groundwork for that argument by showing how a century earlier, the mission and reputation of the First-Year Composition Program and the University of Tennessee drove writing assessment, leaning on grades/grading to reinforce standardized American English as correct and maintain an institutional reputation of prestige.

I do this by exploring archival data on the origin of grades and grading practices in the FYC program at UTK from AY 1926-1927 through AY 1937-1938, a ten-year span. In particular I focus on gradebooks kept during this time period by one of the most important members of the English Department, the namesake of the university library and author of the *Harbarce* Handbook, Dr. John C. Hodges. (I discuss Hodges' biography and its relevance for this project in greater detail below.) The gradebooks record both assignment and final grades, along with qualitative notes on progress, for students enrolled in his courses. I also draw on material that documents course requirements and the college and university mission, to better situate my analysis of grades and grading within the wider institutional ecology. Analyzing the gradebooks, specifically anonymized grade distributions year over year, I show how assessment practices in first-year composition at the University of Tennessee functioned primarily to enculturate students into standardized forms of English, as a Platonic ideal of good writing. Ultimately, assessment in first-year composition, and indeed the course itself, serves a gatekeeping function, sorting students perceived to be effective writers, based on their use of institutionally-approved standardized language, from those who were not. This chapter, then, by highlighting the assumptions embedded in grading, that students are bad or mediocre writers who fall short of the university's rigorous standards, establishes the need for the turn towards antiracist forms of assessment a century later, covered in the subsequent two chapters. My key argument here is that historically, conventional grades and grading practices — measuring students' work against a standard perceived to be or constructed as the singular correct way to write — emerges out of institutional needs and priorities and prioritizes these over students' learning.

My analysis begins with 1926, a watershed year for the University of Tennessee because it marked the move from a semester schedule to a quarterly, a decision that would hold for more than fifty years, until 1984 (Creekmore, 2018, "Academic Calendar."). The First-Year Composition Program also changed its curriculum, adopting a three-course model (ENGL 111, 112, and 113)[31] that called for one writing course during each term of the first year. Intended to cover a student's entire first year, the multi-term model includes skills from sentence- and paragraph-level writing (ENGL 111) to library research skills (ENGL 112) and multiple modes

[31] Evidence of the change in curriculum is clear in Hodges's gradebooks. As early as 1924-1925, the courses Hodges teaches are listed not as ENGL 111, 112, or 113 but as English 1 and English 2. Additionally, the gradebooks prior to AY 1926-1927 list English 1 and English 2 over a semesterly system, further evidence of the change. See "Multiple gradebooks."

of writing (ENGL 113), while also — especially in light of the syllabus used for the course, the *Manual of Instructions for Freshmen Composition* — stressing correct grammar as key to effective writing. While the library keeps forty years' worth of gradebooks, as John C. Hodges taught at the University into the early 1960s, before his death in 1967 (Creekmore, 2018, "John C. Hodges"), I cap my analysis after ten years, in AY 1937-1938. A decade provides a sufficient sample for accurate analysis of grades and gradebooks and their relationship with the early first-year composition curricula[32].

I rely on archival material from the English Department (the *Manual of Instructions for Freshmen English*, 1926) and from the University (the catalog for AY 1926-1927[33], reports from the College of Liberal Arts, and presidential speeches (Morgan, 1923-1938, "The Solution of Present-Day Problems," "The State University in its Public Relations") from the time period to contextualize my analysis. Assessment does not exist independently of the institution but rather is shaped by its larger ecological web of values, goals, and social context (Inoue, 2015). Analyzing these materials altogether gives me a holistic sense not only of how assessment was supposed to work but also how it *did* work, altering student and teacher experiences of writing education and the value(s) placed on particular kinds of writing. All data is publicly available through the Special Collections at the University of Tennessee, listed in the catalog ("Multiple gradebooks," 1921-1962), and was shared with me by university archivists or discovered through archival research, over the course of my research into the history of the English Department, the First-Year Composition program, and its grading practices specifically. I rely primarily on quantitative, big-picture analysis of the gradebooks, along with qualitative analysis of marginal comments in the gradebooks and the additional administrative documents I use. This mixed-methods approach allows me to visualize patterns of grades and grading across the decade of first-year composition courses I analyze, while still situating them within a larger constellation of values and beliefs about the purpose of education as framed at a public, state university in the Appalachian South during the 1920s and 1930s.

As subsequent chapters describe *changes* to grading within the University's first-year writing program, this chapter will lay a baseline for how grades and grading originated within this particular program, allowing readers to see more clearly the changing norms and values encoded in assessment practices. I additionally hope that the "thick description" (Nordstrom 2015, pg. 108) of this chapter, tracing not only grade distributions but also their connection with a specific vision for the university and the first-year writing program particularly, may shed light on other contexts, in particular the origins of composition generally. Despite ample histories of our field (Berlin, 1987, Brereton, 1996, Connors, 1997 *Composition-Rhetoric*), including histories centered on un(der)represented stories, such as Black (Royster and Williams, 1999) and Chicanx (Ruiz, 2016) composition scholars, there remains little work on the history of *grades* in university composition courses specifically; the work that does exist largely takes a birds-eye

[32] Additionally, as Hodges's tenure at the University of Tennessee, Knoxville grew longer, he increasingly taught upper-level courses; later gradebooks record fewer sections of first-year composition courses (ENGL 111, 112, or 113), making analysis of these documents shaky ("Multiple gradebooks"). The larger number of courses taught during his first decade of teaching the new, quarterly English Composition curriculum ensures a larger sample size for stronger analysis.

[33] Cited as the University of Tennessee Record, 1926-1927.

view (Boyd, 1998, Schneider and Hutt, 2014). A case study of assessment at a particular university can help us see in sharper detail the role assessment choices play in shaping students' and teachers' expectations for what "good writing" looks like and what writing education is meant to do — laying the groundwork for the argument I make later in this project, that teachers' expectations about good writing and writing education work backwards, affecting the outcomes of their assessment methods. This study also allows me to take a deep dive into a key archival collection in the UT Libraries, educational records preserved by John C. Hodges. I describe Hodges's significance in more detail below; here, it's worth noting that because of Hodges's prominence in the Department, as its future Head and the author of the *Harbrace Handbook,* the gradebooks I access allows for a particularly sharp insight into the ways that assessment choices are driven by departmental leadership's assumptions about (good) writing and writing education — assumptions that continue, often unconsciously, to guide teacing and assessment. Hodges' status also provides a valuable point of contrast with the contingent faculty who, a century on, exchanged common grading prctices for ungrading. By taking up a locally-specific study of the history of grades and grading in first-year composition, then, this chapter departs from other histories of composition (Berlin, 1987, Brereton, 1996, Connors, *Composition-Rhetoric*, 1997 etc) which maintain a broad, multi-institutional scope, allowing for sharper contrasts between historical assessment practices at a particular university and the push against conventional assessment and emerging, liberatory alternatives a century later (2020-2022).

I open this chapter with a brief discussion of my method(s) in this chapter, then turn to a short history of first-year composition and of grades generally. Then, delving into archival materials, I discuss the University of Tennessee's institutional and departmental mission during the 1920s and 1930s, and I provide an overview of course requirements for first-year composition (ENGL 111, 112, and 113) at the time (AY 1926-1927), particularly its emphasis on grammatical correctness. With this context established, I turn to the heart of this chapter: assessment records kept by John C. Hodges, including final quarterly grades for each student and qualitative notes on some students' progress in the course ("Multiple gradebooks," 1921-1962). I analyze grade distributions from AY 1926-1927 through AY 1937-1938, along with briefly discussing both individual students' progress and Hodges's qualitative comments. This work opens space to discuss the mental model of writing that historical assessment practices push forward, one oriented towards a Platonic ideal of good writing — a term I deploy to gesture towards the ways that grades assume a singular, grammatically-correct version of writing that students may approximate but never perfectly reach. Grades further assume that this Platonic ideal, and students' performance of it, may be objectively determined by their instructors. I conclude with a discussion of how analysis of these documents illuminates the influence of institutional mission and reputation over assessment and curricular choices in the early first-year composition program at the University of Tennessee: the commitment to grading as a means of ranking students by perceived deficits, in order to maintain an appearance of academic rigor. In turn, this analysis lays the groundwork to discuss the motivations and goals driving educators in the UTK First-Year Composition Program in the early 2020s, a century later.

A note on framing

As I mentioned in Chapter 1, I discuss the methods specific to each chapter *in* that chapter. Here, I want to briefly outline the framework I use to approach my archival research, specifically my

commitment to approaching archival materials with an eye to their inherently value-laden nature. My research focuses on a predominately white institution — at the time, an entirely white institution, as the University did not accept a Black graduate student until the 1950s and did not accept Black undergraduates until the early 1960s ("African American Hall of Fame"). Because I draw entirely from holdings in the University Special Collection, this work is bound up with official histories of the institution, which assert their own truth and deny the legitimacy and/or existence of other narratives. My hope is to approach this work critically, acknowledging the situatedness of the archival materials I engage. The story archival research tells changes based on where the research is conducted, whose materials are used, and who is doing the telling. I am indebted to the work of decolonial scholars for (re)framing how I think about working within an archive, in particular Tiffany Shellam's and Joanna Cruickshank's (2019) theoretical piece on navigating and analyzing critical archives[34].

Shellam and Cruickshank argue that through archival research, including research outside of institutional holdings, researchers may challenge colonialist narratives and make room for the voices and stories of minoritized people. Key to their argument is the assumption that archives are more than a collection of "fact[s]" waiting for researchers to discover them but are a value-laden, purposeful collection of material designed to tell a particular story, by the people who collected it (pg. 4) — in this chapter, the university faculty, John C. Hodges especially, and subsequent librarians who preserved this work. Ethical archival research asks not only what the researcher has found but also who created the material and who preserved it, what the creator's and preserver's goal was, and why this material, and not other material from the same time, was preserved, each question rife with power dynamics and social location. Layered on top of these questions is a final one, about the identity and social location of the researcher as well: who is doing the research and why?[35]

Engaging with these questions asks us as researchers to more fully acknowledge how narratives, told at a particular place and time to particular people, (re)write our values. A savvy reading of institutional archives may bring to light narratives outside the dominant, official ones, providing a fuller and more just understanding of our own past, and the pasts of people whose experiences intertwined with our own. While institutional archives such as the University's may be used to prop up official narratives and values, they may also be read against the grain of these narratives, with an eye, as Shellam and Cruickshank suggest, to the spaces in between the records preserved

[34] Despite the fact that my own research is not decolonial, I focus on Shellam and Cruickshanks' work because it was key in helping me understand what archival studies were *about*. Tackling an archival research project during coursework, a project that in fact became part of this chapter, I found myself lost as to how to conduct research within archives — and conducted a mini search for archival, methodological texts while sitting right there, at a table in Special Collections. Shellam's and Cruickshanks' work popped up and has influenced my approach ever since.
[35] A good illustration of this approach to archival research, for me at least, is the musical *Hamilton*. The closing song, "Who Lives, Who Dies, Who Tells Your Story?", draws attention to the way that history changes based on who is telling your story and which historical materials they choose to amplify or ignore. Then, by drawing on Alexander Hamilton's story, amplifying his identity as an immigrant, the musical tells the story of the United States as a multiethnic democracy where people of all backgrounds can thrive. Though this story has received criticism in recent years for a too-optimistic view that elides America's founding racism (see Gordon-Reed in *Vox*, 2016, also Meuwese in the *Conversation*, 2020) , it nevertheless exemplifies the power of archival work to alter the stories we tell each other about who we are and move us towards new ends.

by the university. In this way, institutional archives may be used to expand, critique, revise, and offer alternatives to the stories that the institution tells about itself.

In drawing from the University of Tennessee's official archives, documents produced by the institution, its president, and a lead faculty member, I am keenly aware of the potential of these materials to reinscribe institutional values, asserting the official narrative as a singular truth. Yet, since these materials are key to the larger picture of writing assessment at the University, I choose to use them in this research, being thoughtful about how I approach and interpret them. On the one hand, this focus on institutional actors draws important attention to grading and assessment as practices shaped by authority, whether individual or corporate, and by the value(s) held by authority, laying the groundwork for my subsequent argument that ungrading and other alternative forms of assessment may offer instructors a way to enact their own pedagogical values in the classroom, over and sometimes against institutional values. On the other, given that the source of these materials, their embeddedness with university systems of power, constrains their ability to represent student and contingent faculty experiences, I take care to approach them critically[36]. I ask what stories (about the university, about first-year composition, and about student writers enrolled in composition courses) these materials tell, paying attention to who is telling the story: John C. Hodges, the de facto director of first-year composition during the 1920s and 30s and the University, through its preservation of the material[37].

Indeed, John C. Hodges's authoring of the gradebooks I engage in this chapter makes critical analysis importnt, and also a bit tricky. Hodges is, in a sense, the founder of the UTK English Department and certainly key, even a century on, to its current size and strength, including its financial strength. Hodges served as an English professor at the University of Tennessee, Knoxville from 1921 until his retirement, and death, in the late 1960s. During that time he oversaw the substantial expansion and growth of First-Year Composition; though the *Manual for Freshmen Composition,* effectively a syllabus for first-year writing courses published in AY 1926-1927, is unsigned, it is likely that Hodges contributed to its development and the development of first-year composition curricula; and he oversaw the writing and and publication of the *Harbrace Handbook,* a style guide for correct academic English. The proceeds from the *Harbrace* were later used to start the Hodges Fund, today a multimillion dollar endowment that continues to support research and teaching in the English Department — including, importantly, a grant supporting the use of Equitable Assessment in first-year writing (AY 2021-2022). My engagement with historical gradebooks kept by John C. Hodges in this chapter, then, deepens and complicates my work in several ways. First, his leadership role — at the time my analysis concludes, in AY 1937-1938, Hodges is appointed assistant head of the Department — means that I can treat his grading practices as a proxy for grading practices among the first-year composition faculty broadly. Though there would, of course, have been some variation instructor-to-instructor, it does not seem too far a leap to assume that Hodges's priorities and values in grading student work set the tone among other faculty. Hodges's gradebooks are

[36] The actual narrative of students and contingent faculty is, of course, likely lost to time, whether because students and faculty chose not to make them available to the Special Collections, or whether the Special Collections did not take steps to preserve such stories.

[37] The "Hodges, John Cunyus" entry in *Volopedia* records that "soon after his arrival" in the early 1920s, "Hodges assumed responsibility for the freshmen English program."

additionally preserved given his legacy to the university at the time of his death, a legacy which has since then only expanded, largely thanks to the ongoing circulation of the *Harbrace Handbook*.

This brings me to the second reason my use of Hodges's work is complicated, the role that the *Harbrace* — if not the first-ever writing handbook, the "paradigm" for all that follow (Hawhee, 1999, pg. 505) — plays in establishing his legacy at the University and cultivating a particular approach to first-year composition. That Hodges wrote the *Harbrace* at all is surprising, given that in the early 20th century, literary analysis, not composition and rhetoric, was the more prestigious career choice in English departments widely. Hodges's decision to buck the trend and focus on composition, culminating (in part) in his production of the *Harbrace*, no doubt contributed to the growth of the first-year composition program and his own legacy. Had Hodges not bucked the trend towards literary studies, his gradebooks may not have been preserved, the Hodges fund may never have been established, and this chapter may never have been written. Yet the *Harbrace* also underscores, especially given Hodges's leadership in the Department, the central role that the production of good grammar played in the early First-Year Composition Department. Hodges based on the text on his own analysis of perceived errors in first-year student writing and intended it to serve as a guide, for teachers as well as students, to correct academic English — a Platonic ideal, if you wish (Hawhee, 1999, see also Glenn and Gray, 2017, pg. xx). Engaging the *Harbrace* itself is outside the scope of this project, but its eventual publication, perhaps ten or fifteen years after the gradebooks I study here, and insistance on a singular ideal of English, haunts these pages. In analyzing Hodges's grading practices, I both draw on his legacy, through my use of Hodges Fund money for supporting equitable assessment, and dismantle it, drawing to the surface the values and assumptions about what good writing looks like and who is able to produce it[38]. While I acknowledge the importance of Hodges's work in establishing the ongoing strength of the UTK English Department and validating composition as a worthwhile area of research, I also aim to make plain the hidden assumptions that drive his grading practices, and by extension, grading practices across the First-Year Composition Program. Pairing gradebooks kept by star professor John C. Hodges with the contemporary turn towards ungrading among contingent faculty at the same institution, a century on, allows me to tell a unique story about how universities and university programs change — from histories oriented around inequitable notions of rigor and prestige, towards more accurate, holistic practices that work for both students and teachers.

Additionally, given that in archival research, as in case studies broadly, who is doing the research matters, I want to briefly acknowledge that my own years of experience as an educator influence my interpretation of the data in this work. I strive to analyze data fairly; this is scholarship, after all. But the fact that, as a new teacher, I learned to read grade distributions as a sign of my own success or failure of a teacher, as much as students' success or failure (Wilkinson, 2005), means I am drawn to study the grade distributions documented here — and slightly scandalized by their consistent skew towards Cs, Ds, and Fs. I accordingly approach this analysis with an eye to how the early writing program may have interfered with students' writing education. Yet I also

[38] Hawhee specifically calls out Hodges's insistence on marking Appalachian English as incorrect, effectively penalizing many of the University of Tennessee's own students, who hail from Appalachia. (1999, pg. 512).

acknowledge that my pedagogical commitments, particularly my willingness to take responsibility for students' learning, are contingent and time-bound; the scholars and programmatic structures represented here would not have recognized these commitments. For this reason, I attempt to situate my analysis within their *own* commitments, drawing out both tensions and harmonies between stated ideals of writing education and assessment, and practice. I also explain the ideals of writing and writing pedagogy as they existed at the time — in particular the emphasis on a Platonic ideal of "good writing" — and let those ideas speak for themselves. Ultimately, I conclude that in spite of its democratic commitments, early composition programs such as the University of Tennessee's rely on assessment to force adherence to a single Platonic ideal of writing, effectively barring some students from accessing further higher education and the emerging industrial class.

How did *grading* come to be?

Key to understanding the story told in Hodges's gradebooks is a broader story, about how grades and grading came to be. Surprisingly, this story is patchy, as much of the literature assumes the existence of grades or, at most, documents their faults (see Cowan, 2020 for examples of such work). Only a handful of chapters and books trace where grades came from and how they function(ed) in writing classrooms, a gap which perhaps undermines efforts to reimagine assessment in new, more learning-centered ways.

As with first-year writing, grades and grading emerged from the turn in the late 19th century towards the German research model in American universities. Prior to the Civil War, college examinations were rare, occurring only once or twice a year and conducted by the Board of Trustees (Boyd, 1996). Evaluation at the time also encompassed students' moral character; citing Peabody (1888), Schneider and Hutt point out that at "both Yale and Harvard" grades included "whether [students] attended chapel or showed up to class" (2014, p. 204), criteria that do not necessarily correspond with the quality of student work (Close, 2005). In their earliest iteration, grades are as much an enforcement of character as a measurement of academic skill, a function which plays into composition's eventual role in enculturating students into particular ways of being.

As composition shifted towards theme-based models, grades further developed as a way to rank students — by academic skill, but also by perceived innate potential and morality. The German research model demanded more efficient, objective methods for ranking students according to how well they met a perceived standard for academic work, recording the rankings, and communicating their results both within and outside of the institution. Grading systems met this need. Systematic approaches to grading effectively "creat[ed] modern systems for a modern world," making it possible to "track" students who move among educational institutions (Schneider and Hutt 2014, p. 206, see also Boyd, 1998), in much the same way that contemporary universities rely on high school transcripts and/or standardized test scores to assess students' academic performances. Put another way, grades were explicitly intended to help educational institutions work *as systems*, meaning that they functioned as a way for instructors within a given institution to record and transmit information about the quality of student work to the institution, and for institutions to communicate student performance to outside

stakeholders[39]. In the case of first-year composition particular, its rise as course intended to teach standardized English ran parallel to the push for standardized, coherent grading measures in such courses; Boyd notes that composition "became a fundamental part of the general education curriculum at Harvard during the 1880s, the very same decade that the school moved to replace its numerical scale of grading with one based on a five letter grade (A through E) system" (5). Though, as Boyd notes, there is no causal link between the two events, first-year composition and standardized grading — through historical coincidence — grew up together, reflecting universities' need for objectivity and precision in recording students' educational achievements. Standardized grading, in other words, allowed the institution to pinpoint (presumably) exactly how good students did, to transmit the information within the institution and to other institutions, and to enable comparison among students' performances[40].

The standardizing effect of grades, capturing students' academic performance as a snapshot transmissible to other stakeholders, is not all bad, importantly. By making it easier to record information about students' academic performance, grades made possible a large public education system and ensured all students, not only the very wealthy, benefited. Especially in "increasingly massive urban systems" of American education during the late 19th century, the fact that grades acted as a kind of shorthand for more "detailed accounts of every student's abilities" made it possible for "other parts of the system to work" (Schneider and Hutt, 2014, pg. 206). Despite the flaws with grades, then, it's important to be clear that they played an important role in progress towards towards widespread public education.

Still, grades' history hinges on two concerning assumptions: that student worth can be quantitatively, objectively measured through assessment; and that academic performance is a proxy for moral performance. Although grading schemes were introduced as early as the 1880s, they were still in flux in the 1920s and 1930s. Faced with this disarray, reformers put their "efforts [toward] creat[ing] grading schemes as seemingly 'objective' as mental tests were perceived to be, including particular aspects like distribution across a normal curve" (Schneider and Hutt, 2014, pg 213, citing Cronbach 1975).[41] Increasingly, grades assigned in a particular course were expected to fall along the same distribution as the results of standardized testing, the now-familiar bell curve, so that in any given course, students were quite literally ranked against each other, in comparison to a standard used for measurement. (I demonstrate this ranking in the second half of the chapter, showing the persistence — and deliberate use — of the bell curve within first-year courses in the UTK Composition Program.) Yet as Norbert Elliot explains in *On a Scale*, standardized testing took for granted that literacy skills could be quantitatively

[39] The role of grades in sustaining systems, in part by rendering information in easily-measurable formats, is an idea developed out of a conversation with Jordan Baker, through the use of texts such as *Seeing Like a State* (1999). We planned and started a joint publication but did not finish. See also Elizabeth Williams (they/them), "Do Grades Make Our Lives Worse?" (2022) for an overview of this argument.

[40] The recency of grades also means that when the dataset discussed in this chapter begins, 1926, grades themselves are relatively new, only about forty years old. The emergence and lack of agreement around what grades are and how they should be used is, I think, reflected in some of the early data in particular.

[41] Much of this history occurs in the context of K-12 education, not exclusively higher education. I include it here for two reasons. One, there is overlap between how grades work in K-12 and higher educational contexts; and two, histories of grades and grading are scarce enough that this would be a poorly sourced chapter indeed if I only relied on histories related to higher education.

measured; and that literacy stood in as a proxy for general intelligence (2005, p. 66-67). Taking a (perceived) measure of students' literacy, schools justified decisions about who to admit and who not to admit, who to spend money on and who not to spend money on. Rooted in models of literacy that, especially in the origins of large-scale testing, characterized minoritized populations such as Black or Jewish test-takers, as lacking the perceived intelligence of white folks, these tests contributed to barring minoritized populations, such as Black or Jewish students, from higher education (Elliot 2005, p. 69-73)[42]. That grades and grading are linked to explicitly eugenicist ends underscores the problematic nature of the assumption that educators can put a number to student intelligence and capacity, then use that number to make decisions about students' future.

This problem is made worse by the fact that even as examinations standardized, their connection to moral judgements stuck. Especially in English Composition, grades were often explicitly seen as a marker for students' moral fibre. Indeed, one goal of the freshmen composition course at Harvard was to hone students' self-discipline and character; as discussed above, grades' role was to enforce and affirm the desired character traits. Moreover, grades' ability to fulfill this role hinged on their perception as an objective measure of student performance. Major state universities in the early 1900s relied on the "current-traditional system of grading" (Boyd 1998, pg. 9), evaluating student work and ranking students along criteria based almost entirely on mechanics, how well student work aligned with expectations for "correct" English at the time. At the University of Illinois, Boyd notes, the perceived objectivity of grades, especially grades which centered mechanical accuracy, enabled "instructors to clearly and confidently mark student performance" and thus to "conclusively differentiate among students" (pg. 12). The clarity of this "order and difference" (Boyd p. 13) proved seductive. Grades spoke not only to students' writing abilities but their ability to fit in with and uphold dominant moral and social norms in a systematized, industrializing economy. Students whose language broke with the norm were perceived as a challenge to the status quo, so much so that low grades in English could trigger university discipline. As Boyd notes, the University of California "criminalize[d]" (pg. 10) bad writing, requiring students who repeatedly turned in subpar work to stand before a disciplinary board. Grades, in this way, served a critical disciplinary function in English Composition, binding notions of standardized or "correct" English to moral norms and requiring that students meet those norms through their language use. In this way, grades function to reinforce institutional reputation and maintain a singular standard for good writing, demanding that students either adopt this standard — or leave the institution.

How did first-year composition come to be?

Paralleling the rise of grades and grading is the rise of first-year composition. The story of first-year writing instruction is often told with reference to the prestigious colleges where (in part) it

[42] Elliot (2016) writes that the test-makers later nuanced their claims that literacy tests exposed the perceived intelligence of various racial groups. However, this does not change the fact that large-scale standardized testing was originally grounded in eugenicist assumptions about the intellectual inferiority of some groups and played a significant part of the "admissions policy" (pg. 73) at prestigious universities. Thus, within higher educational contexts, large-scale standardized testing has from the beginning been complicit in discriminating against minoritized populations, barring them from the lofty aims of democracy and citizenry that educational leaders spoke of in outlining their dreams for the university.

developed, Harvard and Yale. Though composition's growth at the University of Tennessee reflects these schools, and not, say, the Normal Schools or other educational institutions designed for students from nondominant communities (Ruiz, 2016), it also makes concrete and particular the origins of first-year composition in the state schools where they continue to flourish, rather than the elite spaces commonly centered in historical narratives.[43] At the University of Tennessee, First-Year Composition took a recognizable shape — the shape it would hold for forty years — in the 1920s and 30s. Reviewing the history of composition, especially within the UTK contexts, I provide a more nuanced sense of the educational landscape, within which the assessment choice(s) reflected in the archival literature may be best understood.

First year composition emerged, at least in part, out of the turn towards public education, a turn reflected in the sense of mission and vision at the University of Tennessee during that period. While, prior to the Civil War, universities largely focused on training a small handful of wealthy students for prestigious careers in law or medicine, after the Civil War, universities shifted. The German research model pushed universities away from transmitting received knowledge towards constructing new, scientific knowledge (Laats, 2018); land grant colleges emerged (Brereton, 1996), and universities began admitting far more students, especially from non-elite backgrounds, to train them for careers in the growing industrial sector (Connors, *Composition-Rhetoric*, 1997, see also Berlin, 1987). Yet as Robert Connors observes, many of these new enrollees lacked "correctness in writing" (p. 9) as universities defined it, and they consistently failed the entrance examinations designed to ensure they were prepared for their upper-division coursework. First-year composition was intended to solve the problem of students' lack of preparedness for university-level writing. Harvard added "a temporary course in remedial writing instruction" (Connors, 1997, p. 11) — English A — and the rest of the country followed suit. The course was not a prestigious one; as Sharon Crowley records, the central place of literary studies in the English Department meant that composition courses were often assigned to "instructors without dissertations or [to] new PhDs" (1998, p. 126), then overfilled with students, with literally thousands assigned to "staff[s] of fifteen" or so instructors," as for instance at Harvard (Crowley, 1998, p. 58). But its existence gave the university a mechanism for ensuring that students developed the grammar, style, and language usage deemed necessary for them to succeed in upper-division courses and avoid "embarrassing themselves in print" (Berlin, 1987). While first-year composition remained discriminatory, as I describe below, its origins are undeniably linked to the increased availability of a college education to the emerging middle class. Far more people could access higher education, and writing courses provided a stepping-stone along the way.

This democratic spirit is amplified at the University of Tennessee, where its status as a land grant university funded to serve the larger community plays into its sense of a public mission. In the 1920s and 30s, the period documented in this chapter, the University of Tennessee was a university on the rise, with a keen sense of its public responsibilities and a hunger for growth. Led from 1919 through 1934 by President HA Morgan, the institution increasingly framed its identity in terms of an obligation to produce research that served the state and its regional

[43] Indeed, the top schools in rhetoric and composition writing programs are never the Ivies, such as Harvard or Yale; they are consistently large, public state universities such as Michigan State University, Purdue University, or Texas Tech University.

industries. Speaking at a convention of land grant colleges, Morgan argues that such universities, including the University of Tennessee, play a key role "in the swing of the pendulum from a conception of education for individual promotion and satisfaction to that of training a people in the interest of the commonwealth" (Morgan, "The Solution of Present-Day Problems," 1923-1938, pg. 3). Echoing the larger shift within universities towards a new, public-minded purpose, Morgan's assertion suggests a democratic undercurrent to the University of Tennessee's mission. Indeed, in that same speech Morgan insists that "it is the glory of the land-grant college that it makes for democracy, serves all the people, and conserves and develops every national resource" (pg. 3).[44] While Morgan, by training an agricultural scientist, emphasizes particularly the value of university research to state industry, such as farming or forestry, speeches like this position the university as bearing a new sense of responsibility for the wider community. Indeed, Morgan specifically speaks to the role that education plays in (re)shaping the landscape in more humane ways:

> Commonwealths are the dwelling places of humanity. Education's challenge is to make them more habitable. The individual is the agent of the commonwealth, and since he is, the more thorough the training he receives, the more ideally that acquired training is administered in behalf of the commonwealth, the more democratic is the republic (pg. 4).

Recalling contemporary emphases on education, and especially assessment, as ecological (Inoue), Morgan's framing of the state as the "dwelling place of humanity" — to be (re)built and (re)arranged via education — is key, stressing at once the interconnectedness of education, and educational institutions such as the university, with the surrounding environment, and the obligation that individuals have to the environment(s) which they inhabit. In contrast with earlier forms of education, this approach is outward focused, (trans)forming the communities to which the university belongs by educating its citizens, and sending them back again into those same communities. This emphasis is reflected in other speeches Morgan delivered at the time, such as a Commencement speech stressing the role of education both in "provid[ing] youth with . . . sound democratic ideals, moral principles, ideals of duty and responsibility, respect for law, service to the state, [and] fondness for home" (5) and also in a leadership role within the wider community (pg. 6, "The State University and Its Public Relations"). As with the history of first-year composition, then, the history of the University of Tennessee is firmly rooted in an ideal public, democratic service; the university claims that it exists for democratic purposes, situated within a particular community and ready to receive students from that community, educate them, and send them back into their communities prepared for civic duties.

This ideal, however, does not align with reality. Indeed, first-year composition as a whole remained a force for gatekeeping and discrimination within the university, relying on emerging

[44] Morgan's lofty rhetoric about the University of Tennessee serving "all the people" conflicts, of course, with the University's historical theft from and exclusion of Indigenous people. The University sits on Indigenous land, part of which (the Ag Campus) was purchased with money — $5 million dollars, adjusted for inflation — from the Morrill Act, which came from selling Indigenous lands west of the Missisippi. Despite the importance of Indigenous land and wealth to the University of Tennessee, as of 2018 only 63 of its more than 28,000 students are Native American or Alaskan. See Ahtone and Lee et al (2020) on "Land Grab" universities, including the page on UTK, for more information.

notions of a standardized English to prevent certain populations from attending or succeeding at college, and to enforce homogeneity among those who actually *did* attend. As Sharon Crowley records, at the time English A got its start at Harvard, standardized English did not yet exist, at least not as we know it today. Rather, A.S. Hill, motivated by class anxieties and the desire to validate English as a worthy, rigorous discipline alongside scientific or classical fields, worked to "define English as a language from which its native speakers were alienated" (Crowley 1998, pg. 60), establishing rules about what was, and what was not, good English. These rules in turn not only served to justify English's position among other scientific disciplines, they also ensured that only a certain *kind* of person graduated from college, the kind of person who could duplicate these rules in their own writing. If this seems anti-democratic, it was. As Crowley notes, Hill complained about the influence of "democratic practices [on] the 'ubiquity of bad English'," saying that if "every man is as good as every other man, every man's English is accounted as good as every other man's" (pg. 62). Hill's comment suggests that if the equality of every human being is fully realized, then distinctions between language perceived as "good" and "bad" will collapse. Language, and specifically English, served as a key marker for social distinctions. This attitude prevailed even at public universities. At the University of Minnesota, Chancellor George MacLean argued for the role of language instruction in homogenizing immigrants to the ways of writing and being expected in the United States:

> So large a proportion of the population consists of foreigners who are ambitious and capable, the University must be content to do a part of this drill. A boy may lead his class in mathematics and Latin and chemistry and still be unable to free his tongue from the Scandinavian accent, or his written page from foreign idioms. . . . The fundamental work of the University must be a struggle for correctness (pg. 17-18, qtd Elliot).

Although MacLean refers broadly to "the University," the attention he pays to both oral and written language centers English, specifically the work of first-year composition. Envisioned as "free[ing]" students from the accents and idioms unique to their home country, composition instruction is meant, here, to ensure foreigners lose their foreign identity, melting into American society via "correctness" in their ways of speaking and writing. On the surface, then, early composition courses may *seem* democratic, and the various university presidents, chancellors, and professors may see those courses are crucial in ensuring students are able to access higher education. But this access comes at a cost. Centering standardized (or standardizing) language forms, first-year composition ensures that graduates maintain dominant ways of writing (and being) and those who do not, or will not, conform are excluded.

First-year composition relied heavily on graded themes to enforce these ways of writing and being. Reviewed carefully by instructors, themes provided an opportunity to, as Crowley notes (1998, pg. 74), discipline students whose writing did not (yet) align with standard expectations (see also Brereton, 1996)[45]. Grades and grading functioned as the chief mechanism for this discipline. As Boyd notes, the "current-traditional system of grading" in use at the time had an "obsession with mechanical correctness at the sentence level" (9). Far more than simply ensuring

[45] Johnson (2021) provides a good overview of how gradebooks generally — including in contemporary writing courses — act to discipline students into particular ways of being; and his work influenced this chapter.

that students did not, as Berlin puts it, "embarrass themselves in print," all the (metaphorical) red ink spilled over these themes pushed students towards adopting the then-emerging standardized American English. Grading of these themes picked up the moral overtone so common in early grades and grading practices. A.S. Hill chalked students' ability to perform well on the themes up to "self-control" (qtd Boyd, 1998, pg. 9), while grades and grading were understood as a form of "discipline" (Tieje, qtd Boyd, 1998). This is key, as it suggests that within the landscape of early first-year composition, students were judged as morally lacking if they failed to adopt the particular ways of writing and being then held as "correct." First-year composition may have played a key role in improving access to college during the early twentieth century, yet the fact remains that the course *also* pressed upon students the emerging singular, standardized (language) norms linked with white, American middle- or upper-class identities, framing those who fell short as morally lacking.

The concern about maintaining certain class distinctions shows up in the University of Tennessee's sense of mission around that time, in spite of all Morgan's rhetoric about benefiting the larger commonwealth. A Biennial Report (1938) put out by the College of Liberal Arts describes its own mission as uncompromisingly idealistic and rigorous: "The University of Tennessee consists in the students and faculty in pursuit of an ideal. That ideal is scholarship." Further, the report adds that "a human ideal must be high" (1). On its face, this wish for the best possible research and teaching is good, pushing the University towards accurate, reliable work with benefits to the surrounding community. In the context of education, however, especially writing education, the spillover effect of this talk about ideals is to frame students as lacking. Indeed, the College of Liberal Arts goes so far as to urge *replacing* the students, bemoaning the restrictions that being a public university — a university tasked with serving the commonwealth, as in Morgan's speech — places on its ability to simply enroll better students. Writing in its 1936-1937 (?)[46] Annual Report, the College notes, "since, at present, we cannot rigidly select our students, the alternative is to encourage more of the best students within the State to come to the University. Many of the superior high school graduates are being induced to attend the larger institutions of the North and East" (12). Here, the College of Liberal Arts seems to express a wish that it could "rigidly select" students, perhaps even from outside the state, to reach its ideal of rigorous scholarship, an apparent contrast with Morgan's vision of a democratic institution. On the one hand is the University's mission as a land grant university intended for democratic, public service; on the other, the College of Liberal Arts's push for excellence in scholarship, to be achieved, if necessary, by changing the composition of the student body.

In fact, students who did enroll at the University took a series of examinations to ensure that, at least by the then-emerging standards of white, largely male academics, they were maintaining correct language use. Remedial education was required for students who failed the examinations. In its 1936-1937 (?) Report, the College of Liberal Arts complains, "the standards for passing English 111 [the first course in the three-course, first-year composition sequence] definitely need to be raised," according to the Junior English Test results (pg. 7). A writing examination required of students during their third year, the Junior English Test carried a penalty of remedial coursework for failure. According to the English Department's report, cited in the College report

[46] Since I began working on this text, the label rubbed off the spine, making it impossible to confirm the date.

of that same year, most students who failed were "assigned to the several members of the Department for individual instruction," with the few students who bombed the test required to take an additional "quarter's drill in composition" (pg. 23). Crucially, the English Department records that in that same year, the "test was given to 543 students, 140 of whom were found to be deficient" (pg. 23) — in other words, a fail rate of 25%. This fail rate underscores that though the Junior English Test is meant to meet the Department of English's sense of "responsibility for the student's use of effective English" (pg. 23), it is also meant, at some level, to raise the barriers for a college education, ensuring that only students who meet the wished-for, institutionalized ideal are successful. To be fair, the College of Liberal Arts did not *want* students to fail, necessarily. Indeed, in that same report the College advises that though faculty should be careful not to inflate grades, they should also take caution not to fail too many students, urging that "more grades of C, and fewer A's, B's, and F's (in general) should be given" (pg. 15). But the fact remains that interpreting a 25% fail rate on test results as a call to raise the standards for passing a first-year course will inevitably result in some students leaving college prior to receiving a degree. The push towards a very particular ideal within the College of Liberal Arts, then, conflicts with the University's broader mission of public service, (re)fashioning the University into an institution centered on reinforcing its desired status of prestige and rigorous education, often at the expense of students' access to or success in college.

Purpose of first-year writing

I have so far discussed the history of composition and specifically its history at the University of Tennessee broadly, showing that despite democratic ideals, higher education centered values of rigor, elitism, and exclusivity, relying on assessment to maintain these values. I turn now to the history and — importantly — the grading practices employed in the University of Tennessee's first-year composition courses, showing how both together construct these ideals within the educational program. I draw primarily on Professor John C. Hodges's gradebooks from 1926-1927 through 1937-1937 ("Multiple gradebooks"), as well as the 1926-1927 university catalog ("University of Tennessee, Record"), and the 1926-1927 *Manual of Instructions for Freshmen English*, effectively a kind of syllabus for the composition course at the time. In beginning my research on this chapter, I hoped for much more comprehensive inquiry into the history of the writing program at the University of Tennessee and in fact spent time talking with archivists in hopes of identifying material such as programmatic outcomes as they existed in the 1930s or documents detailing how assessment was (re)conceived across multiple iterations of the program. Such materials do not exist, at least not in a way accessible to me within the time span of this project. This gap is the norm rather than the exception in composition research, however, as WPAs and other administrators or educators often act on decisions with little or no concrete documentation (L'Eplattenier and Mastrangelo). Such gaps interfere with the ability of research to tell the full story — in this case, for instance, omitting crucial perspectives such as students' own.

Yet what remains is equally noteworthy: institutional documents and those maintained by its representatives, notably Hodges himself. What the archives tell is not a comprehensive story but a more particular one: the institutional role of first-year composition, and of grades within first-year composition, as Hodges and other English Department leaders envisioned it during the early years of the program at the University of Tennessee. As such, the story may help us think more

clearly about the constitutive role grades play in the programs and universities we inhabit, and their relationship with mission and power. While only a short window of time is represented in this chapter, the ten years from Autumn 1926 through Spring 1937, the snapshot sheds light on how grades and grading were originally conceived in the program and provides context for the shift towards ungrading, nearly a century later.

According to the 1926-1927 catalog, the University of Tennessee was at the time experiencing a period of growth. Pointing to "the 1917 Act of the Legislature" and specifically a million-dollar bond "for the erection of buildings and improvements," the catalog frames this time as one of exciting growth for the university, an "infusion of new life" (pg. 36). This growth makes the University "the principal educational institution" in the state, and especially in conjunction with Morgan's hope that ongoing research could have a real benefit for the state, positions it as (at least geographically) an elite institution. This status likely contributes to the enrollment requirements in force at the time. Students in the 1920s were asked to meet prerequisites such as graduating from an "accredited school" (58) where they took "a carefully graded course in oral and written composition," with "instruction in the practical essentials of grammar" (pg. 60). Notably, this course is also required to maintain a "constant insistence upon the elimination of such elementary errors as personal speech-defects, foreign accent, and obscure enunciation" (pg. 60). Echoing the Minnesota chancellor MacLean's concern that universities help immigrants or speakers of other languages "free [their] tongue" from undesirable accents, this requirement assumes a standardized version of English, produced by able-bodied people speaking it as their first language. The catalog, then, lays the groundwork for the singular, ideal version of English subsequently enforced through first-year composition courses and their grading systems, as later sections of this chapter show.

Both the catalog description and syllabus for English Composition center students' production of correct English as preeminent. The catalog records three separate first-year composition courses, one for each quarter (the University of Tennessee shifted to a quarter system starting in 1926, see *Volopedia*, "Academic Calendar.") Each course was rigorously academic, meant to train students in grammar, varied modes and genres of writing, and the use of library resources, among other skills, as the course descriptions in Table 2.1 show:

Based on the catalog descriptions, this year-long sequence of courses focuses on the modes of writing ("argumentation, description, narration") — already on their way out as early as the 1940s (Connors, 1981, pg. 40). As it stands, however, the multi-term course of study described here seems to center training students in skills considered fundamental to effective writing and increasingly moves towards more advanced work, including some research writing (the "library-drill" in ENGL 112). In fact, the courses on offer in 1926 seem to include skills beyond grammatical drills. While ENGL 111, true, prioritizes "sentential analysis, punctuation, [and] outlining," ENGL 112 expands to include "writing . . . of paragraphs," while English 113 delves into varied modes of writing such as "argumentation, description, [and] narration." Whether these course descriptions actually align with the content of the course is another question, but they at least gesture towards a theory of composition oriented towards the modes (see Connors, 1981, 448), perhaps as contemporary writing programs are oriented towards rhetorical choices. The First-Year Composition Program at least on the surface seems oriented towards equipping

Table 2.1: Course Descriptions for ENGL 111, 112, and 113 in the University of Tennessee Catalog (1926-1927)

Course	Description
111	“Sentential analysis, punctuation, outlining. Emphasis placed on the writing and correction of short expository themes. Required of all Freshmen. L.A. First quarter, repeated during the second and third quarters. Students must secure credit for English 111 before beginning English 112 or 113.” (142)
112	”Paragraphing, library-drill, advanced exposition. Emphasis placed on the writing and correction of paragraphs and of a long expository theme. Required of all Freshmen who have received credit for English 111. L.A. Second quarter. Three hours.” (143)
113	Argumentation, description, narration. Emphasis placed on the writing and correction of short briefs and of short descriptive and narrative themes. Required of all Freshmen who have received credit for English 111. L.A. Third quarter” (143)

students with skills intended to prepare them for effective communication in their future courses and professions.

Yet the syllabus belies this more expansive version of first-year composition courses, suggesting that despite attention given to skills such as the modes, the courses emphasize production of a singular, correct English[47]. This goal comes through clearly in the *Manual of Instructions for Freshmen English* (1926). A document which maps out the requirements for each 100-level English course, along with grading expectations, models for revision, and more, the *Manual* functions as a syllabus. Indeed, the *Manual* is uncredited, which frames it as an institutional document, establishing what the university thinks writing should be, or even what "good writing" inherently *is*, rather than how instructors teach or students perceive writing. This is key, as (a la Foucault) university institutions assert their dominance over and against alternative, external, or challenging discourses. As an extension of institutional discourses, the syllabus's approach to grading does not simply say something about grading; it says something about the way that grading functions within academic institutions in ways that affect what language students are expected to produce — and what language is penalized or disciplined, to ensure continuity with larger institutional norms and goals.

In particular, the sections on grading and late penalties in the *Manual* suggest a model of writing defined by adherence to the emerging grammatical norms we know today as Standardized American English. Good writing, according to the *Manual*, is mechanically correct writing. Describing the grading process, the *Manual* notes:

> The student's grade for each quarter will be determined by:
> 1. The quality of his written work and the care with which he revises or corrects the written work after it has been returned to him by the instructor.
> 2. His class work and his preparation of assignments other than the written work.
> 3. His standing in the written quizzes and final examination (11).

This description is noteworthy for its brevity and its ambiguity. A grading schema that throws the weight of students' grade primarily on the "quality of his" writing and "the care with which he revises or corrects" that writing leaves a great deal of room for interpretation. Exactly what kind of writing counts as quality? While Hodges clarifies, noting that "the quality of a theme depends even more upon its substance (thought, ideas, etc.) and its structural arrangement than it does upon mere mechanical accuracy (spelling, punctuation, etc.)," this explanation is also ambiguous, leaving individual instructors to determine whether a student's "thoughts, ideas" and selected "structural arrangement[s]" are high quality or low quality, and what characteristics make them so. The *Manual for Instructions* thus settles power onto the instructors and takes it away from students; decisions about what counts as quality, whether a particular theme

[47] This is not to say, of course, that the modes themselves were not part of the envisioned Platonic ideal, especially as they set rigid expectations around what "good writing" in any given situation looked like, without reference to the wider, fluid rhetorical contexts.

demonstrates that quality, and whether to share information about how quality is determined with students is made at instructors' discretion.[48]

As it proceeds, the grading policy complicates claims about the quality of students' ideas. The *Manual*, despite its assertion that the substance and structural arrangement of a piece matter more than grammar, imposes harsh penalties for grammatical errors:

A grade below passing will be given to a short theme or paragraph for any of the following offenses:

1. Misspelling of five or more words.
2. Two incomplete sentences (Greever and Jones, 1).
3. Two sentences run together, or separated only by a comma (Greever and Jones, 18).
4. Three dangling modifiers (Greever and Jones, 21, 22).
5. Three instances of faulty references (Greever and Jones, 20, 33, 34).
6. Any combination, equally serious, of such faults. (11-12)

Such expectations are strict. While most assignments are short, no more than 300 words, in keeping with the brevity of themes often required in composition courses at the time, at least two assignments (the "long brief for an argument" and the "long exposition with outline") are meant to be up to 3000 words each. By the *Manual'*s standards, a six-page paper with one misspelled word a page or with no more than one incomplete sentence every three pages could fail. In applying such penalties, the *Manual* makes the content of the paper subordinate to its appearance, effectively undoing its previous statement that the quality of the paper matters more than mechanics. Mechanics, it appears, trumps all. The primacy of mechanics is revealing, in that it speaks to the larger programmatic values, situating maintenance of certain linguistic norms as primary. Recalling Berlin's suggestion that composition at this point is intended to enculturate students into the middle-class, the emphasis suggests that grammar and mechanics may be seen as the way for students to write their way into respectability, if they can follow standardized English, as it was then emerging in first-year composition. The *Manual* thus links standardized American English with approved ways of being, the careers and social goals at which it expects its own graduates to aim, then (as I discuss below) uses grades to reinforce these norms among the student body.

Report on grade distributions

That English Composition employed assessment practices which at once constructed notions of a singular, Platonic ideal of correct writing and prioritized those notions as a hallmark both of social respectability and institutional rigor is born out through examination of John C. Hodges's

[48] Also noteworthy is that the *Manual* uses the pronoun "his." While "his" at the time functions as a gender-neutral pronoun, in fact according to the university's own history ("1804. First Women Admitted."), women were "regularly admitted" starting in 1893, more than thirty years before the publication of the *Manual*! In this way, the Manual not only subordinates students to their instructors' authority and so to the authority of the institution; it also reinforces the central role of men in the program and by implication, in middle-class university education. Use of he/his pronouns as the generic form of students reinforces, parallel to the expectations of a singular, white American English, that one way of being, writing, and language-making is expected from all students, regardless of their identities, backgrounds, communities, or rhetoricla contexts.,

gradebooks. The Special Collections at the University of Tennessee keeps forty years' worth of John C. Hodges's gradebooks, beginning in the 1920s and running through the early 1960s ("Multiple gradebooks"). Gradebooks are organized by term and year, with fall/winter courses recorded in one book and spring courses of the same academic year in another book.[49] Each book records the following information:

- the courses taught during a particular quarter, which usually include a 300- and 400-level course in addition to 111, 112, and 113;
- the students enrolled in that course;
- the grade students received on each assignment, including themes, quizzes, and exams;
- and the students' quarterly grade.

The only books kept are Hodges's gradebooks, despite the fact that according to the 1926-1927 Catalog, at least eleven other people taught first-year composition: Johnston, Stone, Coffey, McCleary, Harding, Farrar, Foley, Bryan, Sayre, Green, and Coe (see *University of Tennessee Record*, 142). As I argued above, however, Hodges's authority in the Department makes his gradebooks a workable proxy for how assessment may have worked broadly in first-year composition courses at the University of Tennessee. Hodges comes to set the tone for what the Department — and indeed, composition courses across the United States, anywhere the *Harbrace* is used — considers to be good writing, so by extension, examining Hodges's documents forms a more comprehensive picture of the program's purpose and values, relative to larger institutional goals.

Yet at the same time, I want to be clear that Hodges's gradebooks alone are retained means that whatever story about assessment they tell, it is a story with inevitable gaps, often around how assessment was practiced and received by the most contingent stakeholders, faculty without PhDs and students. Without access to gradebooks kept by instructors (say) like Johnstone, Stone, and Coffey, it is impossible to tell whether Hodges's work represents an institutional ideal or a more widespread practice; similarly, without access to students' perception of their own work and the grades they received, it is likewise impossible to tell how writing assessment impacted their experience in the course or growth as writers. Hodges's gradebooks can speak to how assessment fulfilled institutional goals for writing education, but it cannot tell us how students received that assessment or how other instructors conducted assessment in line with their own goals. Put another way, the gradebooks tell, not how assessment did function within the program, but how the institution intended assessment to function: as a way to ensure that students produced correct English and sort better writers from worse, the better to assert its own social prestige and maintain white linguistic norms.

In the following pages, I analyze the story that Hodges's gradebooks tell about the institutional importance of grades, focusing on aggregate grade distributions year-over-year and qualitative comments Hodges left in the margins. I discovered the gradebooks in Special Collections in March 2020 — in fact, spending five hours on Friday the 13th with my phone, documenting page after page of material, barely bothering to read them, before the university closed to in-person access for months — and continued to work with the materials over the next few years without

[49] Hodges also kept gradebooks for the summer terms, but I have omitted them from collection and analysis.

being required to have any special permissions. Still, though the assessment data is more than a century old, I present it largely in the aggregate, to ensure as much privacy as possible, using a pseudonym when needed, to discuss individual records. I transferred the grade data into Google Sheets and used that to track patterns from one academic year to the next. (Please see Appendix F for a more detailed description of my methods and audit of the quantitative data involved in this chapter.)

My analysis includes grade distributions for the sections of ENGL 111, 112, and 112 that Hodges taught between AY 1926-27 (when the new curriculum was introduced) and AY 1936-37. The gradebook relies on the familiar A-F grading scale, with a few exceptions. First is the *E* grade. The catalog records that E was used to indicate "students whose work is unsatisfactory, but who in the judgment of the instructor are capable of removing the deficiency without repeating the work in class" (43). As the catalog subsequently goes on to describe the process and timeline required for "removing the deficiency" (44), *E* grades were apparently altered subsequent to reporting, perhaps similar to the use of Incomplete today.[50] However, as the following records show, E is used rarely — in some years, it does not appear at all — and dwindles out over the course of ten years recorded here, perhaps replaced by a more consistent use of Incompletes or Fs to indicate those students "whose work is wholly unsatisfactory . . . and shall be required to repeat the work in class" (43).

*D*s are not listed among the failing grades and seem to have permitted students to progress towards the next course in the sequence, making it the lowest passing grade. Importantly, *Fs* are often given to students who complete many, or every, assignment, suggesting that *Fs* do not mark students who do not participate or engage in the course but students whose work is deemed to not be at a passing level. As the catalog records, each grade corresponded to a particular number of credit units, with As earning the most (three credit units) and Fs the least (zero credit units) (pg. 44); the credit units in turn affected graduation and standing at the university (pg. 45), so lower or higher grades played a significant role in students' progress towards degree.

Finally, the College of Liberal Arts' yearly report in 1936-37 lumps three grades together: E, I, and X; in the distributions below, *E* appears separately from *I* and *X*, both of which are represented under the "Other" category. The College of Liberal Arts also lists an *LC* grade, or "left class," which seems to be the early 20th century version of a W. This grade does not appear in Hodges's gradebooks and is omitted. I discovered the yearly report more than twelve months after I initially discovered the gradebooks, when it was too late to redo my distributions without substantial work; what is most important in the tables below are the distributions of the familiar letter grades: A, B, C, D, and F.

The tables below, then, document the distributions of grades assigned in ENGL 111, 112, and 113 over the course of an academic year, typically across two sections for each course. (When Hodges only teaches one section, that is noted alongside the table.) *Perct* indicates the

[50] The 1926 – 1927 catalog also describes the use of Incomplete. In fact, Hodges himself used the Incomplete grade several times, as the tables in this chapter (Tables 2.2, 2.3, and 2.4) show. However, *Incomplete* and *E* indicate different things, and according to the catalog, the process and timeline required for removing each one differs (44). No Es appear to have been recorded for 1927 – 27.

percentage of a given letter grade out of the total (e.g. the percentage of students who earned an A, for instance), while *Perct 111* indicates the percentage of a given letter grade out of ENGL 111 in particular. (I chose to focus on ENGL 111 as the course most focused on sentence- and paragraph-level choices, meaning choices that affect students' performance of Standardized American English; because the course is the first students take in the sequence, the course also provides a clearer snapshot of how Hodges and other UTK faculty engage with students who are, presumably, just beginning their study of college writing.) The additional calculations for 111 allow further comparison with the distribution for ENGL 111 recorded in the 1936-1937 College of Liberal Arts yearbook.

Results

As the table — and associated histogram, added because the visual lends greater clarity — above shows, students enrolled in a 100-level English course were nearly twice as likely to fail the course (12 students) as to receive an A (7 students). C is the most common grade, with 47 students, with D as the next most common (24 students). Historians of composition note that at the time composition courses became more widespread in colleges and universities (the late 19th and early 20th centuries) examinations frequently demonstrated students' perceived lack of writing ability, often failing large proportions of the test-takers (Elliot 2016, p. 11-30). This belief likely was connected to students' backgrounds, as given the enormous uptick in college enrollment at the time, many students likely hailed from families who had never before sent a child to college (Brereton, 1996, Connors, 1997). Seen in this context, the prevalence of Cs and Ds in 100-level courses at the University of Tennessee constructs students, from the institutional perspective, as mediocre writers at best. At the same time, the rarity of As, particularly given the College of Liberal Arts's subsequent emphasis on striving towards an "ideal," makes it possible for the institution to frame itself as a rigorous, elite school.

Table 2.2 documents the number of students who received each grade (A, B, etc.) in any given term, broken out by course (ENGL 111, 112, and 113). At the bottom of Table 2.2, I also provide the *n* of any given letter grade as a percentage of the total number of grades over the course of the academic year (Perct); in AY 1926-1927, 5.8% of the grades given were As. I also provide the *n* of students enrolled in ENGL 111 who received each grade as a percentage of the total (Perct 111); in AY 1926-1927, 7% of students received an A. I have not provided percentages for the other courses for the sake of time/space. Additionally, I did not record Es for 1926-1927. While the E grade was in use at the time, it was very rare, and it is classed under "Other" for the purposes of this project. Es appear as a distinct grade (separate column) in other years.

Alternatively, consider Figure 2.1. Figure 2.1 depicts the number of students who received each later grade as a histogram, showing how grades awarded in ENGL 111, 112, and 113 follow a consistent bell curve. Notice also that there are more Ds given than Bs, and more Fs than As. The predominance of lower grades is striking, as it points towards the narrative that will emerge over the next decade, that students consistently fall short of established standards for good writing.

These patterns repeat year over year throughout the next decade. Notice, for instance, that in the subsequent academic year, students again receive (far) more Fs than As — more than 3 times the

Table 2.2 Grade distributions, ENGL 111, 112, and 113, AY 1926-1927

	As	Bs	Cs	Ds	Fs	Other
ENGL 111 (41 st's)	3	6	14	11	5	2
Perct 111	.073	.146	.341	.269	.122	.049
ENGL 112	2	7	19	6	2	4
ENGL 113	2	9	14	7	5	2
Total	7	22	47	24	12	8
Perct	.058	.183	.392	.2	.1	.067

120 students total

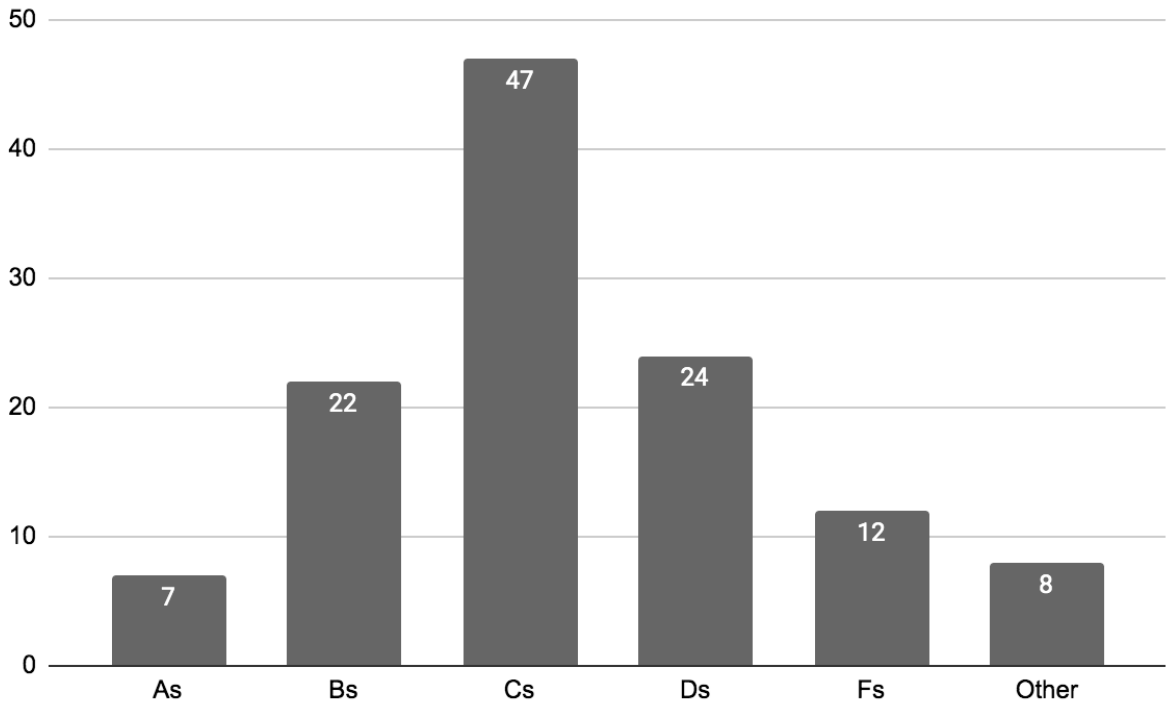

Figure 2.1 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1926-1927

number of Fs! — while Cs remain the most common grade awarded. See Table 2.3 for specific grades:

The predominance of Cs produces, as the Figure 2.2 histogram demonstrates, a neat bell curve, slightly skewed towards the lower grades, underscoring that within the First-Year Composition Program at the time, students are, or are more likely to be, deficient as writers, compared with the standard(s) for good writing established by the grade. (Please note that I have included E to the left of F, given that Es are a slightly higher grade and do not require students to retake the course. Es are very rare, however, and the Figure 2.2 histogram still clearly depicts the bell curve.)

To better illustrate this pattern, of grades falling along a bell curve leaning towards the lower grades, reaffirming students' inability to meet set standards for good writing, I include the Figure 2.3 histogram, which again demonstrates that C is the most common grade, with Ds (in green) typically more common than Bs and Fs (in orange) often more common than As (in blue). I have omitted Es and "Other" for clarity.

Note that in Figure 2.3, AY 1935-37 is omitted because John Hodges did not teach ENGL 112 or 113 that year, meaning the year does not provide a useful comparison for analysis.
Hodges seems to have been tough on students, perhaps as a result of applying the rigid standards specified in the *Manual*; his choices reinforce that at least from the university's perspective, composition students are mediocre writers rarely capable of reaching high institutional standards. Yet the distributions alone cannot speak to how individual students performed across all three sections of first-year composition. I wanted to know whether aggregate bell curves were perhaps masking individual students' improvement from one course to the next. To get the ground-level view of the course needed to determine whether students' grades actually improve from course-to-course, or whether grades consistently construct students as average, I analyze below the grades of one student for each academic year.

Students were selected using a random number generator. If the student received an F or an IC in any one of the three-sequence courses, either of which grades causes them not to advance to the next section, I used the random number generator to select a different student; this ensured that for each student, I had a full three terms' worth of data to analyze. However, because students who receive Fs at any point during their first year are omitted from the data, higher-performing students are overrepresented in Table 2.4. Once the student was selected, I checked to ensure that they remained enrolled in Hodges's section for all three terms, then I recorded their grade.

Table 2.4 shows randomly-selected students' grades course-over-course. Student names are pseudonyms, assigned using Social Security Records about the most common names in the 1930s. Table 2.4 does not include 1932-33, as Hodges taught only a single section of 112 that year and the first two randomizations included students who either failed in the first term or were not in his single section of 112; or 1935 – 36, as Hodges taught only 111 that year and so there are no other courses to compare to. I have listed "No student selected" to indicate that a random comparison is not possible in these years.

Table 2.3 Grade distributions, AY 1927-1928

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111 (39 st's)	0	6	12	11	0	8	2
Perct 111	0	.154	.308	.282	0	.205	.051
ENGL 112	2	8	10	13	1	0	4
ENGL 113	1	10	18	7	0	3	0
Total	3	24	40	31	1	11	6
Percent	.026	.207	.345	.267	.009	.094	.052

116 students

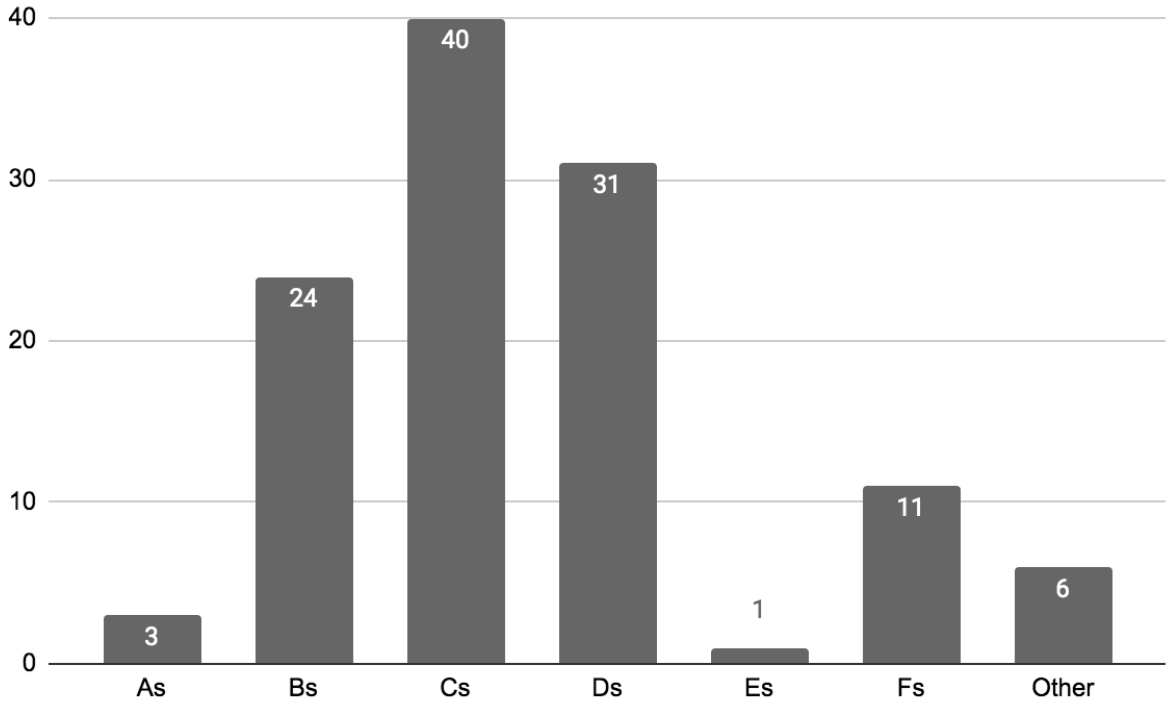

Figure 2.2 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1927-1928

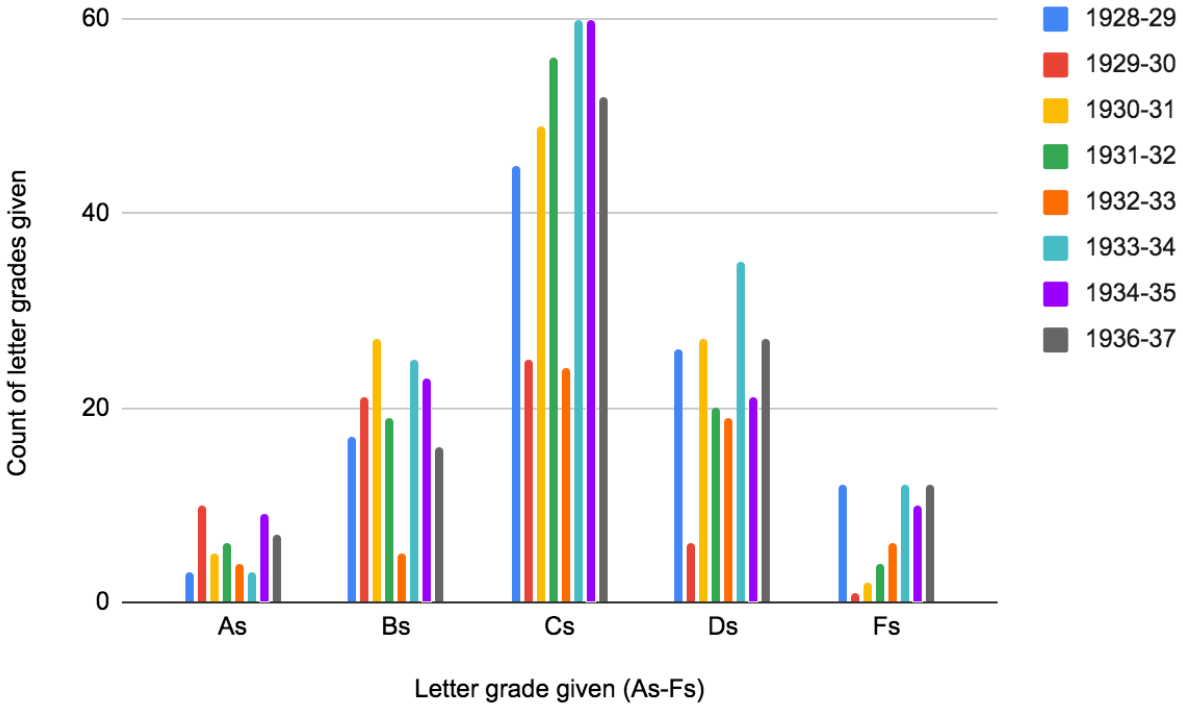

Figure 2.3 Histogram of grade distributions, ENGL 111, 112, and 113, AY 1927-1928 through AY 1936-1937

Table 2.4 Grades received by individual students in ENGL 111, 112, and 113, AYs 1926-1927 through 1936-1937

Year	Student	111	112	113
1926 – 27	Robert	B+	B	B
1927 – 28	James	B	B	B
1928 – 29	Mary	C-	C-	C
1929 – 30	John	C	C+	D
1930 – 31	Betty	B-	C	C+
1931 – 32	Barbara	C+	B-	B
1932 – 33	No student selected			
1933 – 1934	Joan	B-	B	C+
1934 – 35	Patricia	C	C-	C
1935 – 36	No student selected			
1936 – 37	Dorothy	C+	C+	B

In comparing students' grades course-over-course, I wanted to determine whether students were able to work up from the low grades assigned. If students consistently improved their grade, that improvement could signal that the low distributions only recorded students' "starting point," so to speak, with the grades subsequently tracking their learning or growth. That improvement, however, does not consistently happen. Of the nine students included here, only three (Mary, Barbara, and Dorothy) improved their grade from the first to the last class. In at least one case (Mary), the "improvement" is only a third of a letter grade. Additionally, four students (Robert, John, Betty, and Joan) could be said to *worsen* their grade, as it declines by a third of a letter grade or more over the course of the year. The remaining students made the same grade across all the courses. These records may be read in one of two ways. The lack of sustained improvement in students' grades may simply indicate that the courses Hodges taught were growing harder or the genres were shifting, though one would hope that earlier courses in the sequence would provide students the skills to keep up in subsequent ones. The lack of improvement could also indicate that, in keeping with the role of grading at the time to quantify student excellence across the university (Boyd), grading practices at the University of Tennessee primarily functioned to *rank* students and student writing according to a standard, rather than track their improvement, let alone contribute to it.

Analysis

A few things are worth noting about the grade distributions recorded above.

First, and for me the most important, is that these distributions amplify the role of grades/grading in maintaining what I am calling, in this project, a Platonic ideal of good writing. The frequency and rationale for As and Fs in particular suggests that the First-Year Composition Program at the time assumed an idealized, or standardized, notion of "good writing," which could be objectively measured by instructors — in other words, a sort of Platonic ideal, existing in the imagination but never or rarely actually reached by actual students, in actual classrooms. I use the term Platonic ideal deliberately, to capture the assumed singular ideal, the perfect form (Form) of good academic writing against which all students are measured, assume by the persistent bell curve. The ideal of a "pure" or "rigorous" version of English as a standard against which students can be held is of course constructed, in part through the efforts of faculty at elite universities, such as A.S. Hill at Harvard (see Crowley, pg. 60, 1998). Yet Hodges — who orients the *Harbrace* around identifying and correcting 35 distinct errors that students make in first-year writing (Hawhee, 1999, pg. 510) — also participates in the work of constructing the ideal of good writing.[51]

The role of grades, then, is to indicate exactly *how far* students' work falls short, often morally as well as academically. As Table 2.2, Table 2.3, and show, very few students receive an A, usually fewer than 10% of the students enrolled in one of the three first-year composition courses all year. Most students, usually about half of those enrolled over the academic year, receive a C or a D. The distribution follows the guidance from the College of Liberal Arts (1936-1937) that

[51] I do not have time to explore this connection fully (I only located Hawhee's article post-defense, in the final ten days of revising this project before filing with the graduate school) but it is worth drawing attention to the rootedness of this Platonic ideal not only in Hodges's gradebooks but also in his scholarly work.

"more grades of C . . . should be given" (15), hewing to a relatively typical bell curve. Yet in doing so, the grade distribution — especially its emphasis on the C average — highlights a key underlying assumption about the nature of writing education, that "good" — B or A level writing, presumably — is an idealized standard which most, if not all, students approximate without actually reaching, an echo of the College's belief that education is an ideal towards which students and faculty strive endlessly. Students who produce writing below that imagined ideal receive a lower ranking, to designate the perceived gap. A view which contrasts with more contemporary rhetorical practices in which good writing is fluid, determined by the interaction of writer, audience, purpose, and context, assessment against a Platonic standard reinforces emerging notions of a single, standardized English as correct, with other versions of English as incorrect or, flatly, wrong. That the AY 1936-1937 College of Liberal Arts Report, released towards the end of the period represented in this chapter, urges "the standards for passing ENGL 111 . . . be raised" (7) may suggest the increasing stickiness, or persistence, of standardized English as an imagined ideal of good writing.

Second, it is also telling is that *Fs* are more common than *A*s, representing about 10% of the grades assigned in most academic years (1929-30 is a rare exception). Notably, the frequency of Fs *conflicts* with the College's advice in AY 1936-1937 that "fewer . . . F's (in general) should be given" (15), which could imply that in Hodges's courses or other instructors', too *many* Fs were given, perhaps out of a desire to maintain rigor and writing standards. At first glance, the number of students failed is not immediately that noticeable. If 10% of the students enrolled in first-year composition today, assuming a course cap of 20 students, were to fail, that would be 2-3 students per section. What makes the 10% fail rate in Hodges's gradebooks noticeable, is that of the students who receive an F, most do so having submitted much of the work for the course required during the term. This is apparent through analysis of the individual gradebooks, which often show students who ultimately receive an F submitting, and receiving a grade for, regular assignments over the course of the term. In other words, the *Fs* represent what Inoue (2014) calls "quality failure" rather than "labor failure"" students fail because their work was not judged to be of sufficiently high quality to merit a passing score.

The persistence of "quality failure" assumes, and perpetuates, the idea of a single, standardized ideal of good writing that students either meet, or don't, as judged, reliably and objectively, by the instructor[52]. Put another way, the *F* does not say that the student failed to do the work of the course; it says that the student *did* engage in the course but was judged to be a bad writer who produces bad writing — most likely, bad writing as defined by perceived grammatical or stylistic errors. Additionally, because the *F* requires that the student retake the course, it imposes a disciplinary effect for producing bad writing, picking up on the moral component to historical assessment practices. Such effects are institutional and systemic, not personal; Hodges may or may not have thought of retaking the course as discipline, though it is not impossible, given Boyd's (1998) note, described above, that the University of California sent students whose writing fell below the standard before a disciplinary committee. Regardless of Hodges's personal intent, however, the fact remains that persistent Fs especially coupled with a requirement to

[52] Within this model, education is also, at least in part, not (only) about supporting student success but also judging the quality of student writing.

retake the course, frame students enrolled in the First-Year Composition program at the time as both academically and morally deficient, as measured against emerging standards of correct English.

Third, the persistence of a bell curve distribution year over year underscores that within the First-Year Composition program at the time, grades and grading are assumed to be objective and scientific measures, reliably determined by instructors each year. This assumption, of course, conflicts with the reality of actual students, who can (presumably) improve their writing skills and whose abilities are not random in the same way that a coin toss, for instance, is random. In Hodges's gradebooks, however, the bell curve distribution is clear. More than half the class each term receives a C or a D as their quarter grade, with higher and lower grades tailing off at either end. Additionally, the bell curve and relative distributions persist through all ten years, though there are some years (1929-1930) in which Hodges' grading choices vary, becoming easier or harder on students. While course grades are an accumulation of minor grades over the entire term, the persistence of the bell curve in quarterly grades may point to an effort on Hodges's part and on the part of the institution more widely to establish assessment practiced perceived as objective, or that followed patterns considered to be objective. Indeed, that Hodges may have been *intentionally* trying to fit grades to a bell curve is suggested not only by the College of Liberal Arts's (AY 1936-1937) stated preference for C's as a default grade (15) but also by the rapid emergence of quantitative intelligence and literacy testing across the educational sector, as I described in the introduction to this chapter. Indeed, as Elliot notes, these tests, presumed to be an objective measure of "ability" (35) and intelligence, were adopted by such prestigious institutions as the University of Chicago or Princeton (see pg. 101) as a prerequisite for admissions. Educators, wanting to align their own assessment practices with the tests, sought to "create grading schemes as seemingly 'objective' as [literacy tests] were perceived to be, including particular aspects like distribution across a normal curve" (Schneider and Hutt, 2014, pg. 213, citing Cronbach 1975), glossing over the reality that human intelligence and (writing) abilities are not static, random, measurable quantities but varied, multidimensional and under human control. In educational contexts, however, especially literacy testing, bell curves allowed researchers to ignore this reality, prioritizing apparent scientific rigor and reliability to claim objectivity for their own work and determine what groups of people were granted, or denied, access to higher education.

That Hodges's grades (albeit with some yearly variation) appear as a bell curve consistently is significant. Particularly given the College's subsequent direction that teachers award distributions which include more Cs, and fewer As and Fs, essentially a distribution arranged along a bell curve, the persistence of a bell curve in Hodges's own gradebooks may suggest that he, and/or the College, is paying attention to the grade distributions, aiming for a bell curve as implying grades awarded to students as an objective measure of students' innate writing capabilities. A closer examination of the gradebooks shows students' quarterly grades are determined by adding up the grades received on each assignment throughout the quarter, which at once emphasizes the (perceived) objectivity of the process and suggests that perhaps individual grades, as well as term grades, are maintained at a C level.[53] Regardless of whether

[53] I have not actually checked this due to time and overwhelmedness.

individual assignment grades or only final course grades were intentionally maintained along a bell curve, however, the distribution reveals the institution's faulty thinking about grades/grading. Bell curves represent random distributions, yet grades, which are (hypothetically) under students' control and changeable over the course or a term or from term-to-term, are not random. Even within a linguistic framework that values standardized, white forms of English, grades should display greater variation than they do in Hodges's gradebooks, if indeed they recorded students' learning and improvement over the term. No such variation appears. In fact, as analysis of randomly-selected students shows, grades do not improve term-over-term. The persistence of the bell curve, then, especially when laid against students' apparent fail to bring up their grade in multiple first-year writing courses, suggests that grades/grading reduce students and student writing to a measurable construct, not recording learning but ranking by perceived, innate ability and suitability for subsequent education.

At the same time, maintenance of the bell curve throughout Hodges's courses is inconsistent, especially when compared with the grading distributions in college-wide reports. In its 1936-37 report, the College of Liberal Arts published the following distribution of grades assigned in ENGL 111 (pg. 17). The numbers in Table 2.5 are not raw but percentages, e.g. of the students enrolled in ENGL 111 this year, 3% received an A.[54] (Additional percentages were also listed for ENGL 211 and ENGL 311, neither relevant here.)

To compare the distributions of grades in Hodges's ENGL 111 courses to the distribution above, I calculated the percentage of each grade that Hodges gave to his students, omitting the LC grade. As Figure 2.3 above demonstrates, along with the tables in Appendix G, the distributions vary, as in some years Hodges gives As to 0% of the enrolled students, while in others, 18% of the students receive an A. While Hodges maintains (as I discuss above) a bell curve overall, the variation in what that bell curve looks like presents a departure from the College report above, and even from in distributions of Hodges's grades across the entire year. To put this another way, while the grade distributions in Hodges's courses, or in ENGL 111 at the whole-college level, consistently display a bell curve, associated at the time with scientific objectivity, the curve in ENGL 111, as graded by Hodges, often is right- or left-tailed compared with an ideal bell curve.

The skew is telling. On the one hand, the difference here is explained in part by the fact that Hodges's ENGL 111 courses quite simply have a smaller *n* than his courses for the year and a much smaller *n* than all students enrolled in ENGL 111 for 1936-1937, and as the sample size grows larger, statistical abnormalities smooth out. Yet the difference may also suggest that University educators and administrators such as Hodges experienced difficulty mapping their idea of assessment as maintaining an ideal(ized) standard of "good writing" onto reality. While the College of Liberal Arts may *want* to believe that writing assessment can accurately measure students' work, in reality instructors such as Hodges give grades along skewed distributions, consistently more Fs (or occasionally, more As) than anticipated or desired. It is clear to me that Hodges consistently aims at giving grades along a bell curve; it is equally clear that the bell

[54] In the College's report, the table (Table 2.5) also records that 3 percent of students received an LC. Because this grade does not appear in Hodges's gradebooks, I have omitted it from Table 2.5, and the percentages consequently add up to 97%.

Table 2.5 Grade distribution of all ENGL 111 sections, AY 1935-1936, College of Liberal Arts Report, 1936-1937

	Students	A	B	C	D	EIX	F
ENGL 111	2143	3	20	34	24	3	13

curve distributions recorded in his gradebook do not align with the "ideal" curve described the College of Liberal Arts, a juxtaposition which suggests that despite notions of grading as an objective, scientific measure, it doesn't actually work that way. The ideal of a single, standardized English against which student writing is measured conflicts with the reality of actual students, in actual classrooms, taught by actual instructors, resulting in (more) varied grading distributions, even when constrained along a bell curve, as in Hodges's classes. Conceived of as an objective, scientific measure of the messy, human behavior of writing, grades and grading wind up forcing students, and records of student work, into a ill-fitting narrative that serves the university's purposes. Maintaining grading distributions along a bell curve, regardless of students' real efforts to improve their writing or grow, the university props up a reputation for rigor and prestige — at the expense of students, especially those whose writing habits or ways of being do not align with white language practices.

Qualitative data

Just as quantitative measures of student work are designed to discipline students, or bring their writing practices into alignment with the (perceived) ideal of standardized American English. Qualitative measures do the same thing, enforcing a narrow construct of good writing that emphasizes style and, importantly, moral character. A few gradebooks, in addition to recording the letter grade that Hodges assigned to each student, also keep his written assessment of each student, descriptive comments that sum up the students' performance in the course. Comments from 1926-1929 appear verbatim in Table 2.6. For the sake of clarity, I have condensed comments from across all courses and sections of 100-level courses in a given year into a single column. Hodges did not leave comments for the majority of students; thus, the total number of comments does not match the total number of students enrolled in the course.

As comments detailing requirements for students to meet prior to earning credit (e.g. "must repeat English 111 with a passing grade") suggest, the audience for the comments is likely Hodges himself, not the student; they are internal. Interestingly, the use of these qualitative notes slows to a trickle and disappears entirely; in the 1936-37 gradebook, for instance, such comments are missing entirely. All that appears is a list of students' final quarterly grades. Several revealing patterns emerge from an analysis of these themes, including an emphasis on stylistic defects as well as an emphasis on *character* defects in writing. The comments left are frequently critical, and when praise is included, it is often mixed. Notably, a substantial number of the comments (at least five out of fifteen) are critical of students' style. Hodges describes students' sentences as "choppy" twice; "repetitious" once; and in "need [of] variety." Even the comments which praise students' stylistic choices are mixed: "good style — shows promise" compliments students' work while still indicating they fall short of some unnamed standard. Such comments indicate a link between good writing, or the writing valued in the university context, and mechanically correct writing. This link belies the promise given in the *Manual*, that the "quality of a theme depends even more upon its substance (thought, ideas, etc.) and its structural arrangement than it does upon mere mechanical accuracy." Hodges' comments, by camping out on students' stylistic choices, emphasize that whatever points the *Manual* makes about quality as linked to content, in reality "quality" is linked to style and mechanics. Indeed, even comments which don't directly evaluate students' stylistic choices reinforce this view, by describing students' progress in the course as depending on meeting particular standards around style. At

Table 2.6 Qualitative feedback in gradebooks kept by John C. Hodges, 1926-1929

1926 – 1927*	*Very* careless; should do better [italics original] Choppy sentences, — poor transitions Should do B work at least Good style – shows promise Overcome #18 early or go back to 111 Has made decided progress Correct — needs variety style Repetitious Lacks [unclear] Theme 8. Grade not to be given until oral argument is made. Third quarter gradebook includes at least one marginal comment which is written in pencil and unreadable. It includes the word "chapel".
1927 - 1928	Very poor corrections To repeat at least a part of 111 before credit is allowed [on a student who received an E] Copying theme and submitting as his own Or C if RFD wants [on a student who received a D]
1928 – 1929	Faithful work — but [unreadable] To be failed winter quarter if carelessness continued [on a student who received a D- in a course] Careful work, needs more drill Choppy sentence Has made splendid progress, but needs more drill Corrections poor Must repeat English 111 with a passing grade [on a student who received an E]

least two students are described as having "poor" or "very poor" corrections. One student "needs more drill," despite the fact that they have "made splendid progress." Writing is, of course, a skill one improves over a lifetime. Yet that any praise given is consistently contextualized by comments which underscore students' room for improvement reinforces a perception of students as falling short of the ideal of a single, standardized version of writing. Reinforcing the hierarchy between teacher and student, the frame casts students in a deficit role, falling short of institutional standards, as writers.

Crucially, the deficit spills over from students' writing to students' *character*. First-year composition curricula at the time was already explicitly moral, intended, as Sharon Crowley (1998) notes, to "rectify character faults" (pg. 77) and make students "docile" (pg. 78), as they submitted and received feedback on themes addressed required, morally-weighted subjects. The qualitative feedback that Hodges leaves in his gradebook picks up on this moral emphasis. Several comments note students' "carelessness" as writers, a word that conjures up the image of the student as gadfly, ignoring genuine academic responsibilities. Students are not only deficient at writing; their deficiencies in writing speak to greater, moral or personal deficiencies as well. The coupling between perceived deficiencies in writing and in character suggests that at the time, mechanical accuracy in writing was valued not only for its own sake but as a symbol of the kind of person made through an education at the University of Tennessee. Put another way, by extending its field of vision to students' character as well as their writing, assessment frames first-year writing classes as crucial in disciplining students' character for established social roles (Berlin, 1987). Within this framework, falling short of idealized versions of "good writing" becomes a personal failing, so that writing assessment carries greater weight, as an objective measure both of students' academic skill and their character.

Indeed, this connection is often made explicit. In one comment, Hodges links students' perceived character failings to their grade. Speaking of a student who received a D, Hodges notes they are "to be failed winter quarter if carelessness continued." In other words, an *F* grade functions not (only) as a measure of the perceived quality of student work but also as a measure of their moral worth or "carelessness." Students who do not conform to expected academic roles, for instance by demonstrating "carelessness" on their writing, are held back through a low grade. This prevents them from continuing or completing their course of study until they conform more fully with the conservative, institutional expectations. In this way, grades function as a key, defining piece of the larger ecology of the university, enforcing and shaping students into the kind of people — obedient to institutional norms, exemplifying its desired prestigious reputation, upholding white ways of languaging and being — that the university requires.

Discussion

As (part of) a case study, this chapter is limited in its generalizability. Grading practices in the First-Year Composition Program at the University of Tennessee during the 1920s and 1930s are not a proxy for grading practices within other universities at the time, and even Hodges's gradebooks, which *can* tell us something about how his fellow writing professors graded, do not represent the sum total of possible assessment choices in the First-Year Composition Program. Yet of course, as I discuss in Chapter 1, the reality that knowledge of "human action and experience [is] context dependent" (Mishler 2) means only that our claims, or generalizations,

must also be context dependent, rooted in and shaped by the subject: in this case, the interaction(s) between assessment ideologies and pedagogy in the early years of First-Year Composition, as observed at UTK. This chapter plays a key role in constructing such situationally-nuanced claims, since drawing parallels, or contrasts, among multiple sites within the same case makes more precise analysis of the variable(s) possible (see Eisenhardt, 1989, Gomm, 2009, Kennedy, 1979). My hope is that in showing how assessment in the earliest years of the University of Tennessee's writing program was staked on notions of an ideal, scientifically measurable version of English often used as academic gatekeeping will drive home, in subsequent chapters, the value of alternative assessment practices for instructors, at the very same writing program, desiring to empower students and sustain liberatory, expansive conceptions of good writing. Additionally, as I explore the connection between writing program ecologies and ungrading in subsequent chapters, my hope is that the emphasis in this chapter on the important role that the University of Tennessee's mission and purpose, its desire to establish itself as a prestigious university and uphold academic standards, plays in shaping assessment norms clarifies the key role institutional architecture plays in resisting, or inviting, certain configurations of assessment.

This chapter demonstrates that institutional commitments — here, the push for a reputation as a rigorous school — change the shape of writing assessment and sets up a contrast with the different institutional arrangements or architectures, such as the receptiveness or openness to instructors' own autonomy and agency in making assessment decisions for their course, in the same writing program a century later. Since a key argument within this work is that alternative assessment choices such as labor-based grading are not automatically more equitable but bear fruit based on the commitments teachers and programs bring to them, the work this chapter does to chart the commitments that undergird grades' and grading's constitutive role relative to writing education serves an invaluable geographical function, locating subsequent chapters within the larger historical landscape of assessment. To see where we are going with the emergence of ungrading, we first benefit from seeing where we have been: a model of writing education oriented around assessment practices staked on hierarchy, positivism, and a Platonic ideal of writing.

A central landmark within the geography of assessment that this chapter makes visible is grades' function to *rank* students, particularly against a single linguistic ideal. Framings of grading as ranking are pervasive in the literature on (labor-based) contract grading and other forms of alternative grading; as Inoue writes, "grades say little about how or what learning actually took place . . . and *only* offer a hierarchical ranking of the student" (2019, p. 148, italics added). For contemporary educators, however, I suspect this implicit hierarchy, and the associated ranking or sorting into categories, "good" students from "bad," is invisible.[55] Well-intentioned teachers who (still) use grades and grading may not see their assessment practices as ranking. Rather, they may see grades as *earned* by students, not *given* by instructors; as a way to warn students about poor

[55] C. Thi Nguyen has an excellent piece (2022) on the three imagined roles of a teacher: bureaucrat, educator, and cop. Instructors who are particularly invested in the teacher role but still use conventional methods of grading may not be as attuned to bureaucratic and policing roles of grades, where their function to *rank* is most prominent. Certainly as a beginning educator with a literature background, not a composition one, *I* was not attuned to these roles

work and motivate them to improve (Inman and Powell); or, especially for emerging genres such as multimodal work, as indexing a range of rhetorical skills, not a singular ideal (Adsanatham, 2012). Certainly when I first read up on labor-based approaches, I did not see my own grading practice as a crude ranking, arranging students "like laundry on a line" (Potts 2010, pg. 30, citing Elbow, "Ranking"). Grades are undeniably more complex in our highly-mediated age, than in the early 20th century. Nevertheless, what I hope this chapter makes clear is that however complex they are now, grades in the First-Year Composition Program did indeed originate as a hierarchical ranking system. Via grades, leading professors like Hodges signaled exactly how far students fell short of ideals, both ideal (or standardized) writing and ideal behavior patterns; and because the University tracked grades in the aggregate, course-by-course, grades served not only to rank students against the ideal but against each other, creating a permanent record of who excelled — and who did not. Comparison against (an) ideal(s), then, is at the core of grades' function within the University of Tennessee Composition Program, the hub around which its various spokes turn. Even as the Program, like other programs across the country, assess more complicated genres of writing, along more varied criteria, the original purpose of grades to measure and rank students against a presumptive ideal remains, with similar impacts on writing and writing education.

I want to highlight two particular impacts visible in this chapter, and relevant to subsequent work. First is that grades, as a tool for ranking, prioritize the institution and its goals or values above students. By its nature, grading perpetuates a deficit model of students. Year-over-year, Hodges assigns high frequencies of Cs and Ds, often coupling them with comments highlighting perceived defects in student writing and character; these patterns imply that, for Hodges and by extension the institution he represents, students are by and large unable to meet the imagined Platonic ideal of "good writing" maintained by the Composition Program. Students are weighed in the balance and found wanting. Simultaneously, the institution repeatedly emphasizes its own commitments to high ideals, consistently calling for higher standards and lower grades. The resulting effect is that of a see-saw. On the one hand the College, and by extension the English Department, calls for raised standards, in keeping with its own explicit goal to maintain its rigor and (state-wide) elite status; and on the other, students' own achievements remain persistently low, year-over-year. In other words, grading links institutional prestige to students' perceived achievements. The fewer students who receive an A, the more prestigious the University. This relationship is key to this dissertation, as it highlights the way grading puts institutional needs ahead of student or even pedagogical needs; the goals of the institution come first and are predicated on a deficit model of students, especially those who may struggle to observe the standards codified in academic English. In Chapter 4, I explore the co-constitutive relationship between the writing program as an institution and the faculty and students who work there. I lay the groundwork for that study in this chapter, by showing how the historical records, which trace grades and grading distributions traveling up from Hodges' classes to the College, while missives about lowering grades traveling back down, frame evaluative, ranked forms of assessment as fundamentally institutional. Students' own perceived failures reinforces institutional prestige, and the institutional need for prestige locks in a view of students as failing to produce correct English.

Second is that grades' ranking function, manifest across a decade of first-year writing courses, exposes the inherently unreliable nature of grades as a measurement. Presumed to accurately hold students to a single ideal, grades in fact vary wildly, sometimes higher, sometimes lower, year-over-year. Even the bell curve, as I show earlier in this chapter, varies some, with a few years showing noticeably higher numbers of Bs and even As. Some of this variation, of course, can be chalked up to the variable nature of students. Yet the chapter also gestures towards other sources of variation, among them a need to maintain institutional rigor, the tension between the *Manual*'s emphasis on the importance of critical thinking and the high penalties for grammar or stylistic errors in writing, and, finally, the perceived impact of students' character on their grade. Even while the institution emphasizes the importance of maintaining a bell curve in grade distributions, Hodges's own distributions are skewed, occasionally towards As but mostly towards Fs. Grades, in other words, are not a simple, clear measure, tied to a known variable which can be reliably measured. Instead, tied to the notion of standardized (white) American English as a kind of Platonic ideal, grades in fact index a range of overlapping variables, from their own moral character to their use of correct punctuation, ultimately serving as a poor measure for students' learning and growth in the course.. The variation and lack of reliability inherent to grades is important, as, a century later, it is among the most common motivations instructors point to when explaining why they take up ungrading. Hodges's gradebooks document that grades are, in fact, unable to serve as a consistent measure; instructors know this by experience and so, they move away from grades and grading, searching for, if not a better measurement for the *quality* of student writing, at least an assessment practice that will construct a more hopeful English classroom, one which does not tear down but empowers students.

Conclusion

What this chapter shows is that in spite of their longstanding, often unquestioned presence within writing education, grades and grading are neither neutral nor benign. Rather, they (re)shape writing education, orienting it around a Platonic, standard(ized) version of English that signals belonging and acceptance in the middle class. At the same time, as grades are depressed, often in the name of maintaining effective standards, they contribute to a larger institutional image of rigor, supporting claims that the University maintains the very highest standards. The early 20th century was in many ways a transformational, hopeful time for education, as the number of students enrolling in higher education grew exponentially and, for the first time in American history, included those outside the elite classes, destined for careers other than the church, law, or medicine. Many university leaders, including Tennessee's own president, championed this more democratic idea of education, framing the institutional mission as a chance to serve (ostensibly) all students and contribute to the flourishing of the civic community (Morgan, 1923-1938, "The solution of Present-Day Problems"). Grades, however, interfered with this mission, constructing students as falling short of the language standards or character standards (or both) expected of university graduates, ultimately making educational access more difficult. At the same time, grades claimed for themselves a scientific rigor that they could not maintain, leaving the effort to measure the quality of student work (and of students) on shaky ground. Subsequent chapters will show how university institutions, here the culprit in reinforcing the importance of grades as a guarantee of quality and prestige, may rearrange themselves in ways more hospitable to liberatory forms of education. I also explore how teachers, responding to these characteristics of grades as felt inequities, worked within and against institutional or disciplinary norms to take

up assessment practices that centered both instructor and student needs and opened access to writing education.

Chapter 3: Interlude

So far in this project I have charted the origins of grades and grading in the First-Year Composition Program at the University of Tennessee, Knoxville, showing how grades are an institutional device meant to rank students by the (perceived) quality of their work and uphold a university reputation for rigor and objectivity. The next two chapters in this project jump ahead a hundred years, to the grassroots emergence of ungrading from 2020-2022 — a movement which, as I explained earlier in this project, followed in the wake of COVID-19 and the resurgance of Black Lives Matter and attempted to address, if only in part, the perceived urgency for fairness, humanity, and equity in teaching and assessment practices. In the next chapters, I explore how particular programmatic architectures, or structures and relationships that organize the program, solicited or resisted alternative forms of assessment; and I document instructors' experiences with ungrading, showing how the affordances of ungrading are linked to and change with instructors' values, creating a fluid, commitment-driven assessment approach.

In this chapter, which I am poetically calling an *Interlude*, I provide context for those arguments by describing the institutional and social contexts of both the initial turn towards ungrading (2020-2021) and the Equitable Assessment Pilot (2021-2022). This section serves as a partial methods discussion for the two, twin chapters that follow, documenting who was involved in my study, what questions I asked in interviews/focus groups, and how I collected data. I also define how I use the term *ungrading* going forward, situating it in instructors' lived experiences, by drawing on the assessment documents they use in class. I show that UTK first-year composition instructors adopted an expansive model of ungrading that (re)mixed several approaches, including labor-based grading, collaborative assessment, and self-assessment, to decenter quality judgments of student work in favor of fairness, flexibility, and student learning. The definition provided in the second half of this interlude provides context for my subsequent argument, that ungrading functions as a flexible catalyst for instructors to reach key goals around their own agency and equity in the classroom.

While I address the specifics of study population and recruitment in the sections below, I want to say up front that both rounds of this project — AY 2020-2021 and AY 2021-2022 — centered contingent faculty, meaning GTAs and NTT lecturers teaching in the First-Year Composition Program. Certainly, graduate students are not contingent in the same way that NTT lecturers are; they have five relatively guaranteed years of funding, regardless of teaching evaluations, observations, or other metrics outside of their research and progress through the program. Yet GTAs are classed as contingent by the AAUP, pointing to their commonality with NTT lecturers: their locatedness within university architectures, or systems, that exert power over the shape and direction of their work, while giving them minimal (or no) control over those same systems, especially since GTAs and NTT lecturers are by definition not permanent hires ("Background Facts on Contingent Faculty Positions," "Contingent Faculty," AAUP). There are, of course, exceptions. I am one. A GTA, I am nonetheless leaving the First-Year Composition Program at UTK markedly different than I encountered it on my arrival. I would also suggest that thanks to a tradition of nonauthoritarian leadership in the Composition Office, and defined administrative roles for both GTAs and NTT lecturers, the University of Tennessee, Knoxville is better about engaging with contributions from contingent faculty. This reality, however, does not change the

fact that GTAs and NTT lecturers remain contingent; they can and may be dismissed, and have less stake in the program. At the same time, GTAs and NTT lecturers teach, at UTK as at many other research universities, the vast majority of first-year writing courses, as I show in the subsequent chapter (Chapter 4). This reality puts them on the front lines of the assessment decisions I discuss in this project — decisions about how we assess students, what assessment is used for, and how our assessment choices affect the classroom, and the students and teachers who inhabit it.

For me, then, focusing on contingent faculty in this project makes sense for two reasons. First, because contingent faculty teach first-year writing courses most consistently, they are the university employees most perceptive of, and affected by, changes to writing assessment in those courses. Second and related, contingent faculty have less control over their working conditions, including their choices about how and why to assess students. My focus in this project — about the institutional conditions that invite risky, creative decisions about assessment among faculty, and their affective experience of taking up new forms of assessment — answers these concerns, providing writing instructors and administrators a blueprint for thinking through how to cultivate a more effective, fruitful use of ungrading, even among faculty with less institutional power. The work also addresses how to do so in a way that serves not only students but teachers, making space, *contra* their contingent status, for professionally-rewarding assessment practices that answer to faculty's pedagogical goals. I turn now to discussing the specifics of each individual round of the qualitative section of this project, beginning with the AY 2020-2021 round and continuing on to the more formal AY 2021-2022 study.

Institutional and social contexts

AY 2020-2021: Informal growth of ungrading

The story of ungrading at UTK begins early in Summer 2020, at a park on a bright, sunny day, when I met with several fellow UTK educators to talk about using ungrading in first-year composition. Having used a hybrid version of ungrading since Summer 2019, I responded to the acute social crises of that summer by adopting an entirely labor-based approach in my assigned section of ENGL 101. When friends and fellow instructors likewise assigned to teach that summer or fall expressed interest in ungrading, I compiled my ungrading resources — at the time, largely examples of my prior, hybridized approach, along with an explainer and FAQ and links to Inoue's resources — into a Google folder, shared it with those interested, and met them in a safe, outdoors location to describe my approach to ungrading and answer questions. I also offered to address questions one-on-one subsequent to our meeting via Zoom.

Ungrading took off. As COVID and Black Lives Matter had sharpened attention to existing injustices, in academic work as in white American society more broadly, the UTK English Department was, like many at the time, (re)considering its educational and disciplinary commitments through a lens of racial equity and ensuring that faculty members and graduate students had access to equity-related resources and training. I shared the link to my Google folder on the First-Year Composition Canvas webpage. Later that fall, in November 2020, at Dr. Katy Chiles's invitation, I shared secondary research on alternative assessment and linguistic justice at a department workshop. Acting on a suggestion from Dr. Jessi Grieser, I followed the

session up by inviting several participants to talk one-on-one about converting their course to an ungraded approach. We spoke via Zoom prior to the start of the Spring semester.

My plans to research ungrading at UTK were, originally, small in scope. I put together an IRB hoping simply to share in public what my colleagues and I were doing. The more I worked on the project, however, the more interested I became, finally expanding the study into a dissertation.

I collected data between Fall 2020 and Spring 2021. Seven participants enrolled in my study: Lane, Kerry, Jamie, Carter, Taylor, Beckett, and Sterling. Pseudonyms are gender-neutral regardless of participants' real-life gender identity, and participants were invited to choose among several possible pseudonyms. I invited participants to join a focus group or a one-time, one-on-one interview; participants who joined focus groups may have been invited to participate in a further, follow-up interview, depending on my need and their availability.

Our interviews followed a three-part process: 1) an introduction that asked participants broadly about their impressions and experiences with ungrading, 2) a process-oriented section that asked instructors how they developed their ungrading approach and the resources they used, and 3) a conclusion that asked participants about their motivations in using ungrading and any outcomes, including outcomes specifically related to COVID, antiracism, stress, and teacherly identity. (See the interview protocol in Appendix B for more information.) I also asked on the consent form to retain a copy of the materials participants used to discuss ungrading with their class. I use these materials for analysis and context in this project, to ensure that I represent participants' choices accurately and in context.[56]

With the exception of two conference talks in Spring 2021, I began analysis in Fall 2021. Initial analysis included the four initial focus groups or one-time, one-on-one interviews only, as I chose to work with a smaller, more concentrated amount of data to develop my theory before testing and developing it in larger ways. Follow-up interviews are not represented in this study. All data is stored on a password-protected computer or in my UTK G-Drive, also password protected; analysis is conducted in NVIVO.

AY 2021-2022: Piloting ungrading

My own interest in ungrading, along with my colleagues, remained high into AY 2021-2022. In consultation with my advisor, Dr. Jeff Ringer, we decided to offer a small pilot — which we termed the *Equitable Assessment* pilot[57] — to encourage more widespread use of ungrading among first-year composition instructors (ENGL 101, ENGL 102, ENGL 131, and ENGL 132).

[56] While I use the term "participants" in reference to formal research processes, such as conducting interviews and analyzing data, I typically prefer "educators," "instructors," or "teachers" in the findings chapters that follow, highlighting participants/educators' professional expertise, rather than defining them by their role in this study.

[57] As I observed in the Introduction to this project, the term "equitable" in "equitable assessment" was thought more suitable than "antiracist" assessment, given the social and political environment in Tennessee at the time. (Since then, of course, *equity* has itself become a dangerous word.) Thinking about *equitable* assessment also invites attention to the many ways that our assessment practices interact with and support students along multiple facets of their identities (Carillo).

Together with the then-Director of Composition, Dr. Sean Morey, we applied for and received a $6000 grant from the John C. Hodges Fund to support the Pilot, which was scheduled to run for three semesters: Summer 2021, Fall 2021, and Spring 2022. The Pilot, which was not part of my project, sparked more widespread interest in alternative assessment and equity among first-year writing educators, including through a visit from Dr. Asao Inoue in Fall 2021.

Educators who enrolled in the Pilot attended an orientation pre-semester (also open to instructors *not* enrolled in the pilot), received support and check-ins during the semester, and, at the end of the term, shared their ungrading resources with the Composition Office as models for other instructors and received a $200 stipend. While instructors' assessment choices were reviewed at the start of term, the review was conducted with an eye to clarifying and communicating assessment choices effectively for students. In other words, the goal was constructive, not corrective. Instructors had a fairly broad scope in determining how, exactly, they wanted to ungrade, and though most instructors adopted a labor-based model, the specific choices they made within that model varied, as the definitions below demonstrate.

Participants were eligible to join the Pilot for up to two semesters. Participants were recruited from the Pilot as their time in the Pilot came to an end but were of course free to decline to participate. Ten educators overall chose to participate in the research, including four educators from the initial (2020-2021) study; the participants involved in this round of research include the following: Finley, Jamie, Kerry, Oakley, Avery, Sterling, Taylor, River, and Greer. As in AY 2020-2021, pseudonyms are deliberately gender-neutral.

Interviews and focus groups (participants could choose which option they preferred, with participants in a focus group being given the additional option of a one-on-one follow-up interviews) were conducted between December 2021 and August 2022. With the exception of an additional question asking about the intersection between participants' racial identities and their experience ungrading, the interview protocol was largely the same as in AY 2021-2022. I asked follow-up questions when needed to clarify or explore participants' observations. As Chapter 1 notes, I had intended to ask participants to distribute a survey among their students, but for personal reasons, I was unable to do so consistently enough to gather reportable results. For the purposes of this study, interviews and focus groups followed the same protocol as 2020-2021, with Interviews were transcribed by the PI and analyzed beginning in August 2022. Depending on participants' responses and the ebb and flow of the conversation, I occasionally asked participants to clarify or elaborate on a point, but otherwise, the questions were the same.

What do UTK instructors mean by ungrading (2020-2021)?

I use the word *ungrading* to refer to the "constellation" (Schwarz, 2020, p. 21) of alternative assessments that participants used in their courses[58]. For UTK instructors, a version of labor-

[58] As I observed in the Introduction, *ungrading* is a highly contested term, even within rhetoric and composition. Fellow rhetoric and writing scholars have complained that "ungrading" is too "techbro" (Libertz, 2022) and that it obscures the actual practices used (Fernandes, 2022), and that research and writing on ungrading should simply list the assessment choices used. On the one hand, being clear about what alternatives, specifically, we tried is essential for replication of both research and praxis (see Albrecht et al). On the other, asking teachers to list the practice(s) used implies that only one practice is used at once, when in fact practices are often remixed and combined in new

based grading forms the core of this practice, as Table 3.1 indicates. *Ungrading,* however, is more useful as a term as it better captures the array of assessment practices that teachers considered, deployed, remixed, and revised in their course each term. My focus on ungrading as a compilation and reorganization of various assessment practices, importantly, echoes more recent definitions of alternative assessment models such as contract grading as assessment *genres*, composed and recomposed to meet different contexts (Schwarz 2020). Additionally, *ungrading* stresses that participants' approach is not closely modelled after Inoue's — or for that matter, Danielewicz's and Elbow's, or Shor's; rather, educators follow my own approach, adapting it to their own goals as needed.

Table 3.1 lists assessments that educators in the first round of this study (2020-2021) mentioned as part of their ungrading practice. While thorough, the list may omit some information about educators' assessment choices, depending on what they chose to mention and to include in the materials that they shared with me. The left hand side lists the specific practice (e.g. grading matrix that required students to pass the major project), while the right hand side lists the number of participants who included that practice in their own approach to ungrading. So for instance, all seven instructors represented in the first part of this study used labor-based approach, represented in a grading matrix[59]. In this way, the table maps out what ungrading looks like across the instructors who initially took up and deployed ungrading within the first-year writing program at the University of Tennessee, Knoxville. Because my goal here is documented practices shared among instructors, I list only assessment choices that three or more instructors made, which contributes to anonymity[60].

The choices instructors make in pulling together their labor-based approach largely reflects my own ungrading approach, which (as of Fall 2020) tracked the following labor:

- Completion of major projects, along with self-assessments
- Completion of crappy first drafts, revised prior to submission
- Engagement in peer review
- Participation in rubric development discussions
- Completion of homework assignments

---

configurations to reach certain ends. As an umbrella term, "ungrading" provides a handy umbrella term preferable to some alternatives: "equitable assessment" implies that the practice(s) is automatically antiracist, while "alternative assessment" inherently others the practice, defining it against the norm of grading. While I use "alternative assessment" in particular as a synonym, and have used "equitable assessment" to refer to the pilot, I prefer "ungrading" in this project.

[59] By "grading matrix," I mean a table connecting particular forms and quantities of labor to particular grades in the course. Most matrices, including my own, use columns to describe each grade in the course ("For an A you will") and rows to describe the labor called for ("Pass all three major projects. Submit a self-assessment with each." See a sample matrix (my own, for ENGL 102) in the appendix.

[60] Please note that the information in Table 3.1 reflects the fact that some participants, by their own accounting, lean hard on my own Google folder to develop their approach. Accordingly, some of the variation depicted in Table 3.1 may be due to changes in my *own* ungrading practice. For instance, I first adopted collaboratively written rubrics during Fall 2020, meaning that Spring 2021 was really the first time other instructors could use them as a model for their own course.

Table 3.1 Participants' assessment practices, as components of ungrading, 2020-2021

Practice	No. Participants (Listed if 3 or above)
Used a grading matrix/labor-based grading approach	7
Included in grading matrix / labor assessed for the course:	
Pass major projects (some instructors give separate credit for revisions)	7
Crappy / shitty first drafts of major projects	5
Conferences	7
Peer reviews	5
Rubric development sessions	3
Homework / daily work / quizzes	7

- Attendance at one-on-one, teacher-student conferences[61]
- Completion of a pre- and post-semester reflection (the "initial" and "reflective lifeway narratives")

Instructors' use of my own grading matrix as a model is unsurprising. Prior to Summer 2020, I was one of the few instructors in the program using labor-based grading at a course-wide level, though other instructors may have used labor to assess individual assignments[62]. Also, I provided both materials and training for instructors who later opted into the study, via a Google Drive made available in June 2020. As I describe in the programmatic experience chapter, instructors report making substantial use of — or, as Taylor puts it, "basically stealing" — my materials, as a way to boost their confidence in adopting a new assessment method. As I show in subsequent chapters, my materials come to function as key programmatic materials that structure organizational activity, in this case educators' work tracking and documenting student labor (LaFrance, 2019).

That said, I want to be clear that instructors' choices here are *their choices*. Some instructors drew on resources outside my Google folder: Jamie reports "chatt[ing] with a friend" at another institution who "ha[d] been using some style of contract grading for a number of years," and made creative changes to my approach, including prioritizing revisions. Likewise, Kerry describes revising my approach to fit their pedagogical goals: "I looked at your grading matrix, but I pretty much made mine from scratch, just because I think I don't even remember if yours was for what it was for 102 or 101 but it just it felt like I knew what I needed like for an A, for a B, in my class, and so I wanted to make that on my own."

What I want to stress about Table 3.1 is that UTK educators chose to adopt labor-based grading, focusing on students' completion of the work required by the course. This choice contrasts with approaches employed by other users of contract grading, among them Shor, who expresses concern about students "blowing off" (p. 10) course requirements, or even Danielewicz and Elbow, who reserve As for students whose work is perceived to meet A-level standards. To be clear, instructors still paid attention to the (perceived) quality of student work. For the most part, however, instructors did not link their judgment of quality to a student grade; overall, *ungrading* at UTK came to mean a completion- or labor-based system, in which educators linked, or tried to, students' completion of required work with their grade at the end of the course. As Beckett, describing their goal in designing their assessment approach, says: "I always wanted students to pass — I always wanted to be ungrading, I always wanted doing all the work to mean passing the class." While "all the work" varied from instructor to instructor, with some instructors tracking peer reviews and others revisions, instructors nevertheless chose to focus on the *work* completed,

[61] Since Fall 2020, I personally have shifted to crediting assessment via participation (see Gomes et al 2020, along with my approach in Appendix XYZ.) For several reasons, I believe that crediting attendance is slightly different than crediting the labor of writing or preparing to write, and so, I discuss it very little in this project.

[62] In fact, several instructors during 2020-2022 approached ungrading / labor-based grading on an assignment-by-assignment approach, tying letter grades to the work students did *for a project* vs. the work they did over the entire term. Though further research could be done on the outcomes of particular formulations of labor-based grading, my own approach at this time — emphasized by my preference for the term *ungrading* — is that any movement away from conventional, qualitative judgments of student work as captured by a single letter or percentage is worthwhile, and I see this choice as just as much alternative assessment as a course-wide approach.

versus the quality of that work. Crucially, one thing obscured by Table 3.1 is that for most instructors, doing "all the work" resulted in a grade of A. An "A" default is distinctive to the ungrading approach employed at UTK, as other models, including Inoue's, employ a "B" default[63].

Yet while attention shifts away from measuring the quality of student work, the quality measurement does not disappear altogether. One instructor, for instance, reports continuing to use analytic rubrics as a way to provide students "feedback" on "how well [they] achieve[d their] rhetorical purpose for" any given assignment. Other instructors expressed — legitimately, as such concerns continue to circulate even among assessment scholars (Barrios et al 2023) — uncertainty over whether students would be adequately prepared to produce the grammatically-correct language expected in other courses or professional situations. Together, these features gesture towards the way that quality judgments are "sticky" in assessment, often persisting even beyond the adoption of labor-based grading systems.[64]

Ungrading, then, within the context of the 2020-2021 data, comes to mean the practice of connecting students' grade at the end of term not to quality measurements but to the work, or labor, that students do for the course, with the specific nature or content of that work varying from instructor to instructor depending on their goals for the course and the term in which they are teaching. Student success is (re)defined as completing, or passing, the tasks set as part of the course. Key to this definition is that because teachers recognize that "there's only labor involved in writing," as Beckett puts it, by connecting students' grade to the work, or writing and writing-adjacent activities completed for the course, teachers are effectively connecting students' grade to student *learning.* At least in theory, under the ungrading model that instructors employ, students' grades reflect their engagement with and performance of learning, or whether they *do* the things that produce learning. Sidelining questions of quality, this approach centers student engagement with the course materials and with learning.

What do UTK instructors mean by ungrading (2021-2022)?

The emphasis on student labor — to paraphrase the explainer I provide to students, are they "doing the things" — continues in 2021-2022 during the Equitable Assessment Pilot, as Table 3.2 shows. Instructors continued to track largely similar forms of labor, among them revisions, self-assessments, and peer reviews, prioritizing this labor in the grading approach for the course.

A few things are worth observing about Table 3.2[65].

---

[63] The A default likely stems from my own ungrading practice. In my teaching, I prefer an A default as the best option to engage students fully in the work of the course, while also avoiding "busy work" or additional labor, including labor I cannot adequately support during the semester.

[64] This is the subject of my presentation "Whitewashing Contract Grading: How 'Standards Persistence' Undermines Antiracist Assessment," at the *Southern Regional Composition Conference*, 21 April 2021. I showed through analysis of four instructors that those versed in antiracist theory are better positioned to develop ungrading in ways that meet antiracist ends. Instructors who retain elements of more conventional grading practices, such as rubrics, may find they reach antiracist goals in ungrading more slowly.

[65] Please note that "homework/daily work/quizzes" in Table 3.2 refers to the reality that many educators' assessment materials do not distinguish between small, "daily work" type assignments and larger homework assignments that build up towards the larger project. So this category collapses the two possible entries. Educators list these,

Table 3.2 Participants' assessment practices as components of ungrading, 2021-2022.

Components of alternative assessment practice	No. of participants (listed if 3 or above)
Pass major projects	10
Submit a rough draft for each major projects and revise	7
Submit a self-assessment with each major project	5
Conferences w/ instructor	6
Peer reviews	5
Rubric development sessions	4
Homework / daily work / quizzes	10

variously, as "paper prep assignments," "homework assignments," "minor assignments," and "class activities," and more; it was not possible for me to distinguish what, exactly, each kind of assignment was without a more thorough review of participants' teaching documents. Thus, all assignments from daily work up through larger assignments that assist students in making progress on, or partially complete, major assignments are collapsed into the "homework / daily work / quizzes" category in this section

That instructors continue to use labor-based grading, or a version of it, as the primary way they structure ungrading is unsurprising, since, in providing ungrading models to students, the Equitable Assessment Pilot (2021-2022) relied primarily on my own materials and secondarily on materials of instructors who used and adapted my documents. All water flows downwards; to the extent that participants accessed and made use of documents I had originally designed and shared with the department, they worked within a system that connected *ungrading*, broadly, to labor-based approaches. This does not mean that all educators used an identical format since, as Table 3.2 indicates, instructors varied in what labor, specifically, they assessed as part of the approach. Additionally, as with the first round of the study (2020-2021) some educators measured labor on an assignment-by-assignment approach, for instance granting letter grades (As, Bs, etc) to projects, depending on the amount of labor or tasks students fulfilled in completing that assignment. Notably, however, participants center forms of engagement or labor considered key to the writing process: completion of all major assignments, obviously, but also rough or "shitty" first drafts, self-assessments (a form of metacognition), conferences, and peer reviews. What stands out to me about this choice is that collectively, instructors centered students' engagement with the core processes of learning to write, cultivating an assessment ecology that, by its nature, valued pedagogy and growth, making choices about writing assessment aligned with activities valued in writing education (Huot).

Educators made choices about what labor to track and what *not* to track based on their teaching context and preferences. Educators teaching online opted *out* of peer reviews, though they indicated openness to including reviews in the future, which suggests that (dis)comfort with both ungrading and the technological tools we use to conduct our teaching shapes assessment choices. I also suspect instructors' (dis)comfort with the various pieces of ungrading may affect their willingness to use certain components, such as collaborative rubric generation. Though some educators, such as Sterling, center collaborative rubrics, only four educators specifically mention use of this tool. Given the difficulty of collaborative rubric generation, which often requires careful planning and negotiation on the instructor's part, I suspect that educators who are unfamiliar with ungrading or with this practice may opt out of it. That said, some educators, such as Jamie, may include rubric generation but group it under another heading, classing it (for instance) as a form of homework or preparation assignment; it is possible that more instructors than the chart represents included collaborative rubric generation. Instructors may also (re)design their assessment based on their course theme or emphasis — for instance, by allowing students to select from out-of-class activities to raise their grade from an A- to an A. Finally, that participants consistently include smaller assignments in the labor tracked — whatever the name they give to those smaller assignments — suggests the merit or worth of those smaller assignments to instructors, within the scope of the course.

Instructors directly link the importance of students' labor or engagement throughout the course in discussing their ungrading practices, echoing themes in the 2020-2021 cycle. Jamie, speaking of the grade distributions in their course, says, "what I tell students at the beginning of the semester [is] if you put in the work, you get an A. Generally, if you don't put in the work you don't pass the class." While Jamie's comment that students who "don't put in the work", an odd phrasing that for me at least recalls stern teachers chalking a (perceived) lack of quality up to student laziness, obscures it a little bit, notice that the line here, between passing and not, runs

through students' labor, or effort in the course. Students who engage the course pass; students who don't engage the course do not pass. Echoing Inoue's conceptualization of "labor failure," linking *failure* not to students' perceived lack of writing ability but to students' not doing the things, Jamie's response prioritizes student engagement and learning over performance of a particular quality standard.

River, in reflecting on the impact of COVID on their teaching practices, concurs:

> And so moving toward a classroom model like ungrading where again you're empowering students to make decisions for themselves and working within a more flexible framework that still rewards them for putting in the effort but not being held to a particular quality standard that is arbitrarily put in place, by a curriculum, by myself as the instructor. So building in more of that flexibility that lets life happen in small ways that don't even have to be addressed, so removing some of that friction, or the stress for students and the stress for myself about how to accommodate when every day there's a new accommodation, so I really liked the model for the ability to — reorient what is at its core most important, you know?

What is important about River's observation here is that they link the reorientation away from quality judgments towards student effort to accommodating students' human needs and complex lives, a particularly challenging task during COVID. Their comment gestures at the way that labor-based grading is not a sheerly intellectual or pedagogical choice for many instructors but one made, as in fact Chapter 5 explores, deeply intertwined with affect and embodied human needs. By virtue of its centering students' effort and engagement over an arbitrary standard, as River suggests here, ungrading is also perceived among instructors as a practice capable of accommodating the fact that they and their students are not "brains on a stick" (Smith) but embodied people, with real needs. I draw this out because, as I turn to the next section of this chapter, which explores how educators in 2020-2021 *felt* about ungrading, I want to be clear that at least for UTK educators, this practice cannot be defined (solely) by the constituent labor measured, or even the choice to measure labor over perceived quality, as though UTK instructors' experiences were achieved via, or could be reproduced by, a paint-by-numbers process, deploying a set of practices to reach a set of outcomes. *Ungrading*, for educators in 2021-2022 as in 2020-2021, comes to mean, yes, a linking of students' grade with their completion of the various tasks associated with learning in first-year writing courses. But ungrading *also* means a set of practices that acknowledges the whole humanity of student and teacher needs and (re)considers how students and teachers relate to each other in productive ways. I explore this dimension of my definition in the following section.

How do instructors at UTK experience ungrading?

Mapping out the core practices in instructors' ungrading models is crucial but not sufficient. Additional attention to how they *experience* the model — in other words, the kind of work that instructors see ungrading as asking them to do in their classrooms, and who ungrading is asking them to be — is also key in any definitions. Throughout my conversations with instructors, they drew on similar language to highlight ungrading as work that reorients their own sense of identity towards the classroom and students, highlighting the importance of *ungrading* as a non-

authoritarian, liberatory assessment practice in ways that are not fully articulated via the specific design choices documented above. This analysis, for reasons of time, covers the 2020-2021 data only.

To analyze educators' language choices related to how they described and conceived of ungrading, I mapped out the contrasts instructors drew between their experiences with conventional assessment and their new experiences using ungrading and alternative assessment[66]. This focus allows me to track change across time, documenting the shift(s) in instructional roles and behaviors. The lists below document these differences. The first group is listed behaviors or roles associated with conventional grading; the second notes new behaviors or experiences that instructors associate with ungrading.

Grading
- Justifying grades
- Teacher as arbiter or authority
- Punitive relationship with students
- Penalizing students

Ungrading
- Giving feedback to students
- Conversations with students
- Engage with student ideas
- Student growth
- Freedom (for instructor and student)
- Learning, learning environment

As these lists show, UTK instructors experience grading as an institutional practice requiring them to perform highly-regimented quality control, which inevitably situates them in an authoritative role. Kerry and Jamie both characterize grading as forcing them into the role(s) of "intimidating arbiter of grades" or "final arbiter of what counts as quality writing," respectively. As arbiter, deciding which papers earn As and which earn Bs, instructors take on the (unwanted) role of judge or authority, applying institutional standards to determine what counts as "good" writing. This work necessarily also requires that instructors penalize students who fall short of the perceived ideals for writing, which instructors *also* dislike. At least three people (Taylor, Jamie, and Carter) use the word *penalize* directly, insisting that they don't "want to penalize students … from lower income families" (Beckett) or language learners (Carter), while other instructors at least gesture towards the concept. Lane links their "hat[red of] grading" to the work of assigning a number to students, or student writing: "you are an 82, and you are a blah blah blah, and it felt so personal." What Lane's comment underscores is that because grading demands educators measure student work against an institutional ideal, educators often experience grading as measuring the *students themselves* against that ideal, and penalizing those who fall short. It is, for many teachers, a uniquely personal and painful task.

[66] The practice of "values coding" (Saldana 2009) provided inspiration for this work, though I did not actually create a set of codes or enter the results into my codebook.

Perhaps because of its personal nature, grading also pressures instructors to focus feedback on justifying grades. All seven educators, in fact, mentioned that grading required this kind of response. Jamie describes spending "forty or fifty percent of my feedback" explaining "why I docked points from [a] paper" — which suggests that the more time instructors spend justifying grades, the less time they spend helping students learn. What UTK instructors find with grading, in other words, is that it imposes a particular kind of relationship with students and student writing: one geared towards order, hierarchy, and discipline. This is exactly the function that grading held in the early years of the UTK Writing Program, as the archival chapter shows; its recurrence here gestures towards authority, and the (perceived) ideals maintained by that authority, as its core values.

Ungrading, on the other hand, invites UTK instructors to step outside of this institutional dead end. While I have outlined the specific practices it may involve above, I want to emphasize here that instructors experience it as freeing, for both them and in turn their students. No longer forced into the role of arbiter, UTK instructors increasingly find it possible to give constructive feedback on student work and engage in a wider variety of ways with student responses. Indeed, freedom and openness dominate instructors' experiences with ungrading. Carter observes that ungrading gives students "freedom to explore their own writing," Jamie notes the "freedom to engage with my students' ideas rather than justify a grade," and Sterling, concurring, points out that ungrading "opens the door" for new ways of reviewing and responding to student work. In other words, *un*grading is experienced by instructors as a set of practices that center both students and teachers, and the writing experience; what is *actually* judged to be key to good writing development, rather than what the larger bureaucratic institution thinks is key, becomes important. In turn, instructors find that because ungrading frees them from the authoritarian role imposed by grading, it makes space for them to support student learning more fully. Instructors frequently use the language of *growth* and *learning* in talking about their experiences with ungrading, suggesting that this shift towards better feedback giving practices and freer engagement with student writing ultimately contributes to development and flourishing (both academic flourishing and human flourishing) in a way that grading cannot do. For UTK instructors, then, ungrading is defined or experienced as an assessment practice that, oriented to students' academic and personal well-being, puts teachers in the drivers' seat and lets actual learning, rather than institutional norms, drive progress.

Conclusion

Ungrading, then, functions as an umbrella term over the next two chapters, capturing, yes, a set of practices educators used but also educators' affective relationship with that assessment: a way of assessing invested in acknowledging students' effort to learn, their human needs, and their diverse language practices, via crediting their work over the term. *Ungrading* at UTK (2020-2022) is broadly labor-based, with an A or A- default, and though the particular labor teachers value — self-assessments, collaborative rubric generation, revisions, homework, or conference attendance — varies, that teachers prioritize labor points to a desire to free students from arbitrary standards of "good writing." Refocusing on student work — are they "doing the things" called for by the class — as the measure by which students pass the class and secure a grade opens space for teachers to cultivate a more humane, embodied experience in the classroom, and

refashion their own teacherly identity away from "sage authoritarian" towards guide and collaborator. More than, say, "alternative assessments," *ungrading*, despite its contested use in the scholarship and in public social media conversations about writing assessment, gathers these interrelated practices and emotions into a cohesive, purposeful whole. I deploy it in this project to capture the way that, though specific practices varied instructor to instructor, instructors shared a common goal to find a way out of the top-down institutionally-mandated practices that defined grading in the early years of UTK Composition, centering students' work or labor in the course as a way to more equitably determine students' grade and cultivate a liberatory assessment ecology.

UTK instructors' experiences, of course, do not provide *the* definition for ungrading. My goal here is not to write a definitive study, only to develop an internal, cohesive sense of the diverse, labor-based, non-hierarchical approaches that came to be known as *ungrading* within the context of the University of Tennessee, Knoxville. This definition is nonetheless important, both in providing clarity about what "ungrading" means in the context of this project and illustrating how ungrading, far from being a constant classroom-to-classroom and program-to-program, often takes on distinct shapes and goals within the context of particular institutions. As I show in subsequent chapters, educators' exchange of the historic emphasis on measuring students against institutional norms for assessment practices that valued students' effort, collaboration, and reflection has far-reaching impacts across the various institutional ecologies, from the classroom up through the writing program, department, and university. The subsequent chapters explore these impacts, and the context(s) for ongoing assessment change and development at UTK. Educators' strong, felt sense that grading penalizes students as learners but ungrading frees them from arbitrary, restrictive expectations, that grading boxes them as teachers into unwanted identities but ungrading opens up freer ways of being as educators raises important questions about what programmatic structures facilitate the turn towards ungrading, and what commitments other teachers need, to see their ungrading work bear fruit. I take up these questions in the next two chapters, tracing the effects of educators' choice to ungrade and the institutional support that sustained — and in some cases, undermined — this choice. I show in the first (Chapter 4) that, given institutional investment in imposing grading as a way of maintaining the university mission and reputation of rigor, institutional investment in cultivating ungrading — work that extends to deliberately renovating programmatic structures — is crucial for more liberatory assessment practices to flourish. In the second I turn to individual instructors' experiences, showing how the specific values or emphases which educators brought to ungrading affected the outcomes or affordances that it had in their classroom, work that invites readers to reflect on, and alter, their own pedagogical commitments.

Chapter 4: The Role of Hostile vs. Hospitable Programmatic Architecture in Supporting Ungrading

Although this chapter appears midway through the project, it was in some ways the first part to be conceived, the origin for all the others. I first heard of contract grading at the Conference of College Composition and Communication (2019), as Inoue argued in his keynote that alternative forms of assessment could assist in bringing about racial justice in educational spaces . A first-year PhD student committed to finding new and better ways to teach my students, I was intrigued but skeptical. As we poured out of the dark auditorium into the brightly-lit hall, I turned to a colleague and remarked that fascinating as contract grading was, I thought Inoue much better positioned to employ it than us. "He's a tenured professor. We're graduate students."

I overcame my hesitation that summer. Assigned a summer section of ENGL 101, I worked hard to condense its sixteen-week curriculum into the five weeks alloted to summer sections and wanted a way — *any* way — to streamline assessment in the course. The question of how teachers' assessment choices are shaped by, and in turn (re)shape, the wider academic contexts they inhabit remained. My ability to "go rogue," adopting and designing a new assessment practice, was predicated on my own decade plus of classroom teaching experience and a corresponding (over)confidence in my pedagogical ability. Not all teachers are equally ready to jump into unfamiliar practices. The longer I advocated for ungrading at UTK, the more attention I paid to how educators' position in the wider institutional landscapes affected how they perceived, used, and adapted alternative forms of assessment. Other graduate students, contingent faculty, or faculty from minoritized communities — including Black faculty — may find it difficult to take up ungrading, regardless of their personal interest in the approach, given that Black and minoritized faculty often work harder to gain students' respect and feel more vulnerable within the institution (McCloud, 2023, Craig, 2021). This chapter takes up the question of how writing programs resist, or invite, the use of liberatory forms of assessment among contingent or vulnerable faculty, with a view to casting light on considerations for writing programs looking to encourage or develop the use of alternative forms of assessment in their own contexts.

Until very recently, this question is addressed in almost none of the literature on contract grading. As my introduction describes, the focus of ungrading literature is typically on the singular classroom as the domain of teachers, constructed or imagined as a space where decisions may be made largely without regard to the writing program, academic department, and/or larger universities. This focus is, as I discussed, especially prominent in Inoue's work, where he frames assessment ecologies — along with the potential for antiracist action — exclusively in terms of the writing classroom, largely omitting questions about institutional context(s) that, like Russian nesting dolls, give shape to the classroom. Inoue, however, is not alone in this focus; other, earlier researchers, less interested in assessment ecologies, similarly focus on the classroom as the locus of change, taking for granted that educators are in a position to enact change. Danielewicz and Elbow, for instance, urge teachers to take up contract grading as a way to improve students' learning experience, while Shor argues for contract grading as a way to flatten inherent classroom hierarchies and ensure students have access to power and agency, even over and against their teachers' own preferences about how the course is run. Even

research focusing on more than one educator (Inman and Powell 2018) focuses on educators *as individuals*, agentive professionals capable of choosing, or not, to use ungrading, independent of the surrounding ecologies — program, department, and university.

Yet for writing assessment, as I show in the introduction to this project, to be ecological means that even at the classroom level it is inevitably shaped by the surrounding institutional contexts, or — as I describe them in this project, the institutional *architecture*. Institutional ethnography, the primary methodological frame I use in this chapter, holds that standpoint, or the location of an individual within the larger landscape or field of an institution, is crucial to understanding both that individual's activity and the larger activity of that institution (LaFrance, 2019). Who an assessor is, what institution they belong to, and the position they hold within the architecture of that institution, matters for understanding their assessment choices and connecting those choices to our own contexts.[67] Much of work on contract grading is taught by scholars with access to some position of privilege. As of the publication of their signature works on contract grading or alternative assessment, Ira Shor, Peter Elbow, and Asao Inoue were all tenured men; Jane Danielewicz, a tenured woman. In contrast, though we do not, to my knowledge, have concrete data on the *rank* of faculty using contract grading, we do know that labor-based grading is highly associated with first-year writing and professional writing courses (Litterio 2016, 2018) — and that first-year writing courses are most commonly taught by contingent faculty, per the National Census of Writing (University of California San Diego, 2017) and women (ADE Ad Hoc Committee on Staffing, 2008[68]). My point here is not to criticize Inoue, Elbow, and Danielewicz for their positions of power. Far from it! In fact, as I discuss in this chapter, those positions of power likely afforded them the space and security needed to theorize about and practice alternative, liberatory forms of assessment. My point is simply (as I suggest in the introduction) that only by engaging the larger institutional architectures that give shape to the classroom, and reckoning with the ways those architectures shape the classroom and the activity of educators, especially *contingent* educators, we will miss a more robust understanding of how teachers come to use ungrading and their affective experiences and motivations for using it in their classroom.

More recent scholarship[69] begins to fill this gap, addressing how teachers' experiences of labor-based grading or ungrading is affected by their own place within the larger institutional circuitry,

---

[67] To be sure, research has acknowledged the influence of individual *students* on the effects of or shape that assessment takes, often justly so. Inoue writes that the student resistance Spidell and Thelin (2006) see in their research may be explained given students' likely enculturation into a white racial habitus (2015). Less attention has been paid to the influence of *teachers'* locatedness, especially within the institution and institutional discourse(s) about grades, on the outcomes and shape of alternative or equitable forms of assessment.

[68] The ADE Report is, of course, fifteen years' old as of the writing of this dissertation. (Even the Writing Census, from 2017, is more than six years out of date and does not reflect any of the massive changes across higher education prompted by the pandemic.) Nor does this data reflect more contemporary understandings of gender as a spectrum. More recent research, however, does not seem to exist — likely in part because of the difficulty of undertaking a large-scale analysis of the demographic makeup of first-year writing instructors across the United States.

[69] A few research articles (McDonald 1973, McDonald 1975, Purdy 1975) from the 1970s argue for large-scale, programmatic uptake of alternative grading practices. These, however, are pragmatic in focus, centering particular recommendations without contextualizing them in scholarship or persuading the reader in their favor. Additionally, these are effectively lost to the mists of time; they have had very little impact on the current discourse in alternative assessment practices. Accordingly, I pass over them here.

tied into the whole in ways that affect its use, design, and outcomes; this work also grapples with the importance of educators' own place in the institution, whether at the edges or the center, in determining their experience with alternative assessments. Sherri Craig (2021) points out that top-down mandates that require a particular form of assessment, especially as a "solution" for racial inequity, place Black women faculty in particular in a precarious position, forced to justify to students the use of alternative assessment despite the reality of ongoing, unsolved racial inequity and even violence across the university. I remain haunted by Craig's account of how labor-based grading, a practice many white teachers, like me, take up as a way to work towards racial equity, can in fact backfire, making things worse and not better. I want to *remain* haunted by Craig's work, which is echoed in other literature on ungrading and assessment broadly, arguing that Black faculty often are asked to put in more work than their white counterparts to justify ungrading and build relationships with students (McCloud 2023) or observing that Black faculty often take on greater risk from across the university, including from peers and supervisees, regardless of whether they are using conventional or alternative forms of assessment (Perryman-Clark 2016). Crucially, some of these faculty — McCloud in particular — express a great deal of hopefulness about ungrading, arguing that it challenges unhelpful power dynamics between students and faculty and invites alternative, liberatory ways of being in the classroom. Still, their honesty about the challenges of ungrading pushes scholarship more broadly towards that same honesty: How can we reckon with the ways that teachers' decisions about whether and how to use alternative forms of assessment are made within, not independently of, the institutional structures that we call home?

In this chapter, I draw from institutional ethnography to explore these questions about how programmatic architectures resist, or encourage, the growth of ungrading across multiple programmatic ecologies.[70] I describe my use of institutional ethnography fully in my methods chapter. Here, I want only to emphasize that my use of this methodological framing, and key associated terms such as boss texts, invites a bottom-up approach warranted by the literature on ungrading, meaning that I am not primarily interested in what the writing program or department intends by permitting ungrading. Rather, I am interested in how the most contingent and vulnerable members of the program — the GTAs and NTT lecturers I interviewed — experience the architecture of the First-Year Composition Program as hostile or hospitable, resistant or inviting, to their decision to use ungrading in first-year composition courses. By framing the institution as "co-created" (LaFrance pg. 28, 2019) among those who inhabit and work within it, whether along hierarchical or peer-to-peer relationships, institutional ethnography gets at the way that ungrading is not an individual, classroom-level choice that can be made, or not, by instructors independently of programmatic considerations but one that is made (and then re-made each semester) based on teachers' position within and influence over the larger program. Assessment choices are not made solely by individuals, within individual classrooms, but are instead made within a particular professional landscape and bear its stamp. Our choices about assessment look different, institution to institution, thanks to the architectural features, systemic and dynamic, of the programs we inhabit. The educators represented in this chapter make choices specific to the places they inhabit within the First-Year Composition Program, the English

[70] I am grateful to both Lacey Wootton and Stacy Wittstock for talking about institutional ethnography with me. Their explanation of the method(ology) made it possible for me to use and describe it at least semi-accurately, and I am sure my framing here owes something to their discussion of it with me.

Department, and finally the University of Tennessee, Knoxville, and to the relationships with peers and colleagues that define their location within those places. Framing this next section within the methodological lens of IE captures a pattern I noticed in speaking with instructors, that beyond the affordances or outcomes they valued in ungrading, their experiences with alternative assessment are powerfully shaped by the larger institution ecologies.

In documenting the influential relationship between institutional ecologies, especially the ecology of the First-Year Composition Program, and the emergence and growth of alternative forms of assessment within writing courses the University of Tennessee, I theorize hostile — or hospitable — programmatic architecture, or the ways that institutional structures and processes open or close off choices among contingent faculty. I draw the term from urban design, where "hostile architecture" refers broadly to social control of behavior in public spaces, specifically the behavior of certain classes or groups of people. Hostile architecture asks who is allowed (or restricted) from using a space, and what actions are permitted (or denied) in those spaces. Classical examples of hostile architecture include metal spikes along the ground to discourage homeless people from sleeping in public areas, or sloped benches to prevent skateboarding or sleeping (Quinn, 2014). More recently, cities have removed benches from train stations, also to discourage homeless people from sleeping there — even during very cold spells in northern cities such as Toronto (Chellew, 2021). The goal, as James Petty explains (2016), is to "exer[t] some form of disciplinary control over users of public and urban spaces" — especially in "contested spaces" (pg. 69) where multiple users, including people from minoritized and/or vulnerable communities, such as homeless people, want access to the space, for divergent (though not necessarily competing) reasons. By erecting hostile architecture, such as spikes or sloping benches, city leaders exert power over the behavior of these vulnerable communities and often drive them out of the spaces altogether, or frustrate their ability to achieve their values and/or to meet their needs.

In this project, I draw on "hostile architecture" as a metaphor for the architecture or construction of writing programs and other institutional spaces, in ways that invite or discourage certain actions, among them the choice to use more liberatory forms of assessment, by certain groups of people, especially contingent or vulnerable faculty such as GTAs and NTT instructors. This is not to say that TT/T faculty are not also affected by hostile architecture. My project, however, focuses specifically on contingent educators, as they typically hold less power within the institution and may experience more difficulty in renovating hostile architecture. As I explain in the subsequent pages, hostile architecture can refer to any of a number of aspects of writing programs — from work processes such as yearly teaching observations or mentoring relationships, to common or "boss" texts such as syllabus templates or programmatic rubrics for grading student writing. Any feature of the programmatic landscape or architecture plays a key role in inviting — or restricting — instructors' agency in choosing antiracist, liberatory forms of assessment.

One key difference between hostile architecture as a term for urban design and as a term I adapt in this project to describe certain configurations of writing programs or institutions is that in urban areas, hostile architecture is typically (always?) intentional. In writing programs, however, I argue that hostile architecture may be unintentional. Hostile architecture may arise as well-

intentioned, necessary features of the program, as for instance, the mentoring relationships and teaching observations I discuss in this chapter; such features are not intended to hinder ungrading, yet the impact or function of these features is often the same, creating friction or resistance with more innovative assessment practices. The fact that programmatic architecture may be *unintentionally* hostile is key, however, as it invites writing educators and administrators interested in supporting alternative forms of assessment to consider how their program resists the turn towards liberatory assessment — in ways that go far beyond their own interest in such assessment, or desire to sustain it within their professional spaces.

By contrast, I coin the term "hospitable architecture" to refer to the ways that the programmatic structures or organization make space for teachers' decision or agency to use more liberatory forms of assessment in their writing classes. Closely linked to hostile architecture, *hospitable* architecture is distinct in two ways. First and most importantly, hospitable architecture is oriented towards inviting choice and agency, versus discouraging it. The move towards ungrading at UTK was bottom-up, grassroots level, not top-down; instructors opted into ungrading, not out of it. This decision was intentional on the part of those of us working in the Composition Office — intended to avoid conscripting people into work that they felt made them more vulnerable, or did not fit with their professional goals or identities (see Craig, 2021). Hospitable architecture, then, is not the reverse of *hostile* architecture; it is not an attempt at social control, pushing people towards ungrading rather than away from it. What hospitable architecture does, as I describe in this chapter, is open up choice for instructors — to make ungrading feel like a legitimate, safe choice for them, versus one constrained or resisted by the larger infrastructures of the university. This brings me to my second point, which is that hospitable architecture is oriented around ensuring teachers have the resources to choose alternative or antiracist forms of assessment, not cutting those off. Unlike hostile architecture, which constricts behavior, hospitable architecture widens it.

Finally, I use the word "renovate" to capture the processes by which writing programs may rework — expand, rearrange, or dismantle — certain aspects of the institutional space, in order to more fully authorize and invite educators' decision to use ungrading in first-year composition courses. Renovation, or change within, a programmatic architecture is essential to maintaining that space in ways that serve contemporary needs, especially since (as I discuss above) hostile architecture is not intentional but accidentally, as older features of the programmatic space, built to serve good ends, suddenly come to resist more innovative practices. I borrow the term *renovation* as suited to the architectural metaphors of hostile or hospitable architecture. In the Introduction to her collection *Renovating Rhetoric in Christian Tradition,* Elizabeth Vander Lei envisions renovation as a process of expanding or reimagining the boundaries of religious discourse to include new, often more liberatory ways of being (xi). I transfer that definition to the domain of assessment research and extend it here, deploying *renovation* as a metaphor to invite attention to the ways that writing programs and departments may reimagine their own boundaries towards more liberatory ends. The process of renovation, ideally, enables students, teachers, and administrators to lean in, more freely, to new, often risky, assessment practices.

In the following chapter, I point to features of the programmatic landscape as hospitable, among them key relationships with peers and supervisors, ample written resources, ready access to

expertise, and more. In exploring institutional architecture, I see myself as expanding and extending the ecological or environmental metaphor. Architectures, like ecologies, are interconnected and mutually reinforcing; a house would not remain standing if the beams did not support each other. The term architecture additionally allows me to consider how a given program may be simultaneously hostile and hospitable, much as an old house may be newly renovated in some places and yet creak and groan in others. Writing programs, and other institutional structures, are at once more intentional and deliberately constructed than ecologies, harder to change, and likely to include both out-of-date and shiny new features side-by-side, creating unintentional points of resistance or friction. The term architecture allows me to explore hold these realities in tension in the following pages, in exploring how the University of Tennessee First-Year Program's hostility/hospitality towards ungrading and alternative assessment.

This chapter is divided into two sections. The first documents both the hostile and hospitable architectural features of the UTK First-Year Composition Program during 2020-2021, when the program provided only informal support; and second, exploring the deliberately hospitable architectural features in 2021-2022, when the FYC Program, through the English Department, provided intentional support for contingent instructors using alternative forms of assessment. Using key terms and concepts from IE, the first half introduces the characteristics of hostile or hospitable programmatic architecture, focusing on 1) work activities such as regular observations or observational criteria that may actually discourage ungrading, 2) the role of boss texts or organizing documents such as sample grading matrices in supporting ungrading; and finally 3) the potential of relationships among both peers and supervisors to support and organize ungrading choices.[71] I briefly define each term or concept, then turn to educators' specific experiences as framed or clarified by that term, to show how ultimately, ungrading emerges from institutional activities and relationships as well as individual educators' goals and desires[72].

In this first section, I draw on my analysis of focus groups/interviews with the seven instructors who participated in the first round of my research from 2020-2021: Lane, Kerry, Jamie, Carter, Taylor, Beckett, and Sterling. I show how, especially during this initial year of support, their experiences shed light on areas of the Composition Program that proved hospitable, or hostile, to their decision to ungrade. My analysis is context-specific, of course, in keeping with my use of case study, affected by the unique structural features of UTK's First-Year Composition Program. The reality that all FYC courses are taught by contingent faculty, the autonomy that faculty, and the Composition Program, have over curricula and course design, and the semi-regular scheduled reviews, among other features, all play a role in constructing the program as hospitable or

[71] As I discuss in my methods chapter, I do not consider this dissertation or even this chapter to actually *be* an institutional ethnography, in part because I omit key terms from my analysis, such as "ruling relations." While further research may benefit from following this line of thought, using institutional ethnography to frame the data, as it stands, my use of three select terms helps me draw out important contextual factors that explain why ungrading is easier, or harder, for instructors at UTK.

[72] A complete analysis of UTK instructors' use of ungrading through the lens of institutional ethnography is beyond the scope of this dissertation, which primarily uses IE as a heuristic to clarify programmatic or other institutional affordances. Nevertheless, the fact that institutional ethnography centers the dynamic, fluid relationship among instructors and the wider institution makes it a strong fit for this project and further research would benefit from a more extended inquiry.

inhospitable, as I describe below. Institutions and institutional leaders outside of UTK may find this chapter valuable in considering how their own programs may proactively support alternative assessments or simply cultivate a more hospitable architecture for educators to make their own decisions about assessment.

2020-2021 data

As I discussed in the Interlude, AY 2020-2021 provided contingent faculty (GTAs and NTT instructors) with *informal* support for ungrading. Using the term "informal" to draw a contrast with the more formally constituted Equitable Assessment Pilot, funded by the English Department during AY 2021-2022, I explore in this chapter the supports available to instructors via existing programmatic structures — for instance, collegial relationships with peers and supervisors already primed to approve ungrading, models for ungrading within the UTK Composition Program that I readily provided to friends and other interested practitioners. Following my own, largely successful experiments with ungrading, contingent faculty took up the practice and employed it in their classrooms. This chapter, grounded in my analysis of the 2020-2021 data, describes how instructors' experience using ungrading — their sense of whether the practice was viable for them as instructors, given their contingency and vulnerability within the program, and their comfort employing or expanding it in their classrooms — is linked to the informal support they received through the UTK First-Year Composition Program. Crucially, in keeping with IE's focus not only on official, sponsored supports but also on the constitutive, bottom-up activities and relationships that construct organizations in particular ways, I discuss in this first section the informal networks and supports contingent faculty accessed in using ungrading. This allows me to document programmatic hospitality — and hostility — towards the use of alternative assessments in ways that exceed the intention or purpose of the people who actually inhabit and work within the program. Not every feature of the programmatic landscape that supports ungrading is purposeful, and instructors or administrators wanting to think through the potential for antiracist assessment in their writing program should consider how preexisting or unintentional features of its architecture make space for, or restrict, such action. I focus on three features in particular in the following section: boss texts, or the authoritative, organizing documents that ensure consistency in instructors' working habits and choices; work processes, or the ongoing activities that carry out the work or goals of the institutions, and the constitutive peer and hierarchical relationships that structure teachers' work within the writing program.

Boss texts

One key piece of understanding the programmatic architecture as hostile, or hospitable, towards ungrading among contingent faculty is analysis of what La France terms boss texts. By boss texts, LaFrance means the array of documents, forms, and other written materials that describe or chart work done by the institution — for instance "position descriptions, advertisements for employment opportunities, and contracts" (LaFrance, pg. 81, 2019) or other documents that ensure agreement around what the work of the organization looks like and the values that ground it. Mapping out the boundaries or landscape of certain kinds of work within the organization, texts function "architectural[ly]" (LaFrance 81), ensuring consistency in work practices, both among individuals within the institution and between individuals and the larger organizational mission. They are, in essence a "standard[izing]" (LaFrance 43) force, ensuring that all members

of the institution receive the same information and perform the same dance. In the UTK writing program, I suggest that boss texts include documents such as sample assignment directions and syllabus templates, which provide not only boilerplate language about grades/grading and late and attendance policies but also the university-wide grading scale (e.g. A = 93-100%).[73] UTK currently has a First-Year Composition program of approximately 90 instructors and 200 sections of ENGL 101 and ENGL 102. Across such a large program, standardized syllabi and other instructional documents are crucial to provide guidance and consistency.

Yet UTK instructors who take up ungrading describe relying on other, quasi-institutional texts in lieu of conventional boss texts. Though this move provides crucial structure, teachers also find that it sets up conflict with other, established institutional processes, as I explain in the sections on work processes and relationships. Neither the use of alternative texts nor the resulting tension is surprising, given that boss texts make legible the "norms [and] values" characterizing an institution and so invite friction, as the individuals within that institution develop their own set of values (LaFrance, pg. 82, 2019). Put another way, the more individual instructors alter or deepen their professional commitments, a change likely to happen the longer they use ungrading, as I discuss in Chapter 5, the more they find that the boss texts previously available to them no longer work for their needs, and the more they require alternative texts to structure their work. This is certainly the case for UTK instructors using ungrading, who turned to alternatives to standardized syllabuses and grading practices, including the Google folder with my own resources, to structure their choices about alternative assessments. Such quasi boss texts thus function both to make the programmatic architecture more hospitable towards contingent faculty's decision to ungrade — and expose the hostile architecture in the program, given the way standardized assignment or assessment templates discourage the use of ungrading.

Most UTK instructors, according to the focus groups/interviews I conducted, drew on a Google folder I put together with my ungrading resources, as prior authoritative texts became unavailable to them. (Please see Attachment: Contents for Ungrading Folder Circulated at UTK for more information.) I created in the Google folder in Summer 2020 for fellow UTK instructors also teaching that summer or fall to use in designing their own approach. Given that interest in alternative grading mechanisms was high that summer, I made the link public and shared it on social media sites as well (the link has since been restricted to UTK-only; you can find the folder included as Attachment: Contents for Ungrading Folder Circulated at UTK). All seven instructors who participated in the 2020-2021 round of the research report drawing from this Google folder and note its usefulness in helping them get started with ungrading, though some instructors also revised the materials to suit their own teaching preferences and style.[74] Though

[73] A term originally coined in the context of research on public management research, "boss texts" encompass "authoritative" documents that map out work in such a way "that an institutional course of action can follow" (Griffith and Smith 2014, p. 12). Griffith and Smith give as examples "law, policy, [and] managerial objectives" (pg. 12), which suggests a *bossy* dimensions to such texts; these are texts which tell people what to do and, perhaps also, what will happen if they don't. In higher education, where authority is often more distributed and diffuse and autonomy is prized, such bossiness is rare, or at least *rarer.* For this reason, though syllabus templates or assignment directions are largely optional, I list them as "boss texts" since they articulate programmatic objectives, synthesizing the diverse array of composition courses around a single more consistent set of goals or curricula.

[74] Though I met with most instructors prior to their starting ungrading and discussed my approach with them, instructors made choices suited to their own course contexts and teaching style. These choices were not made by me,

some instructors drew on resources outside of my Google folder (Jamie, for instance, reports using articles in academic newspapers and resources supplied by friends at other institutions), my own Google folder seems to play the largest role in shaping instructors' decisions about how to design their own assessment approach. Given its significance to people's choices about ungrading, the Google folder effectively functions as a quasi boss text, organizing and ensuring some consistency — similar, though not identical, choices about assessment — among instructors who use ungrading. That some of the most consistent choices instructors made, such as the strong preference for grading matrices as a way to organize and communicate ungrading in the syllabus, appear in the Google folder points to the folder's important role in ensuring the program was hospitable towards alternative forms of assessment among UTK contingent faculty.

As a quasi boss text, the Google folder both provided models to stand in for conventional boss texts, such as the syllabus's boilerplate language on grading and grading policies. In doing so, it gave people confidence in their work. Instructors point specifically to the value in seeing how ungrading works in their own local context. Perhaps because ungrading remains a diffuse, highly-variable practice, with models that look very different from teacher to teacher and program to program, the context-relevant approaches in my folder made ungrading seem possible. Jamie, noting the importance of a "grading matrix that is custom-made for UT's composition course, says: "In theory I liked contract grading, but I never knew how to adapt it for my specific circumstance until I had seen how you had done it here with the specific assignments and the specific schedules that UT has and so being able to see that has helped me concretize it [and] enabled me to actually put it into practice." The attention Jamie gives to the "specific assignments and the specific schedules that UT has" is key, since it underscores the value of boss texts in coordinating instructional activity with larger university schedules and expectations. Institutions such as writing programs automatically depend on standardized assignments, templates, and rubrics; these direct work not only within the program but also ensure consistency with the larger institution, aligning the program with the university landscape at large (LaFrance 2019, p. 82). Because the Google folder coupled ungrading into the existing curriculum, along with its standardized assignments and schedules, it effectively serves as a new, quasi boss text capable of revising existing texts in ways that open space for ungrading.

The Google folder's function as a quasi boss text for educators, crucially, gave them confidence in their work. They could use it in much the same way as new teachers might rely on a syllabus or prewritten assignment guidelines, to ensure clarity and effectiveness in teaching as they adopted new-to-them pedagogical practices. Taylor, reflecting on the use of the Google folder, says: "I felt like I was on firmer footing using materials that had already been created rather than trying to write it myself when I had never done it before." In other words, as a quasi boss text, a (set of) document(s) capable of regulating instructors' use of ungrading to ensure consistency with the program and use of best practices, the Google folder corrects for instructors' inexperience with new methods of assessment. Instructors outside of UTK who use alternative assessments often struggle with ensuring consistency with programmatic expectations and with

---

and I needed to take time during interviews to ask instructors to describe their grading approach to me and explain how it works. In a few cases, I was surprised by instructors' choices.

knowing whether their assessment approach is effective (Schwarz 2020, p. 65). What Taylor's comment, along with widespread use of the Google folder among participants, indicates is that its status as a text capable of organizing and standardizing work corrects for these common difficulties, allowing instructors to surmount some of the challenges commonly associated with taking up new assessment mechanisms. This structure may, in turn, ensure long-term use of ungrading, as instructors feel on firmer footing in their pedagogical decisions. (Indeed, Taylor participated in the second round of data collection as well, AY 2021-2022, suggesting sufficient confidence in this method to keep up its use.)

Despite the strengths of seeing resources such as the Google folder as kind of boss text, two caveats remain. First, the Google folder is only a quasi boss text. At the time of the 2020-2021 research, ungrading was largely informal within our department, simply another choice people made about their assessment, rather than more concrete support, such as "alternative teacher evaluations" or "stipends and funding" for use of labor-based grading (Schwarz 2020, p. 63). My Google folder was available to instructors and the department at large upon request, and at the invitation of a faculty member, I shared some of my research (secondary research, at the time) on ungrading and antiracist assessment in November 2020. Still, I was a fellow graduate student and the documents carried no official authority[75]. Second and relatedly, because the Google folder organizes people's choices about assessment only in an informal, optional, fluid way, people's decisions to draw from the folder, and in fact their decision to use ungrading more broadly, brought them into tension with other more formal, departmentally sponsored boss texts and work processes in the program that remained geared towards conventional assessment practices. Typical processes such as teaching observations, along with the documents that guide and structure such observations, communicate an expectation that instructors will use conventional letter-grade assessments, putting instructors who use ungrading into a tight spot and constraining the institutional landscape in ways that make the program more hostile to contingent faculty's decision to ungrade. I discuss these difficulties in the next section.

Work processes

"Work," in institutional ethnography, is defined to include not only paid labor but the entire range of tasks associated with a person's role in a larger organization (see LaFrance 2019, pg 40-41). Work processes, on the other hand, are discrete and numerable; they are "how the work gets done" (LaFrance 2019 p. 41, italics omitted). The work of UTK educators represented here, whether graduate instructors or non-tenure-track faculty, is, broadly, to teach first-year writing. The work processes, however, include the many distinct activities which make up that work and ensures it happens according to expectations. Work processes include not only tasks such as the planning and delivery or lessons or the designing of a Canvas site but also participating in annual observations and reviews. Crucially, such processes are often tightly regimented, often through organizational documents such as review guidelines; for lecturers, though not for GTAs, they are tied to promotion. As a facet of the larger institutional ecology, these work processes have the

[75] This, by the way, does not really change in the 2021-2022 data collection, since although the Google folder was made available to people as part of the departmental pilot, it was made available *as an example*, rather than as a template that people were to feel obligated to follow, and in fact some of the instructors included in the pilot research, adopted approaches to alternative assessment substantially different in key ways.

potential to expand or to constrict the organizational landscape in ways that make more or less room for ungrading as a viable assessment option. My research demonstrates that work processes may function as a form of hostile architecture, especially in unintentional ways, among them yearly teaching observations and mentoring relationships.

One example illustrates both the pressure that work processes may exert on ungrading in the classroom, as well as the emerging tension between ungrading as aligned with quasi institutional documents such as a Google folder and the existing institutional documents of syllabuses and review processes. One participant, a GTA undergoing a regular teaching observation, reports that the observer's resistance to view ungrading as a legitimate choice was surprising and confusing. (I omit the GTA's pseudonym here to increase anonymity.) During AY 2020-2021, the GTA was observed and reviewed as part of regular teaching practices at UTK. Sharing the experience with me afterwards, the GTA noted their "surprise" at the observer's apparent disapproval and skepticism of ungrading. Teacher observation processes for graduate instructors at UTK typically ask that final copies of graded papers be submitted for review, showing a range of grades (e.g. A-NF[76]), and teachers are evaluated on their grading practices. In other words, the UTK observation process assumes the use of conventional grading methods. Ungrading is a wrench in the spanner of this work process. As the GTA explained, the observer seemed "confused" by ungrading and asked a lot of questions that the GTA found difficult to answer and experienced as critical; ultimately, the GTA was left wondering whether the evaluation would be "negative or positive," based on the observer's reaction. My goal here is not to critique the individual observer — it's a good thing to ensure that teachers have a rationale for their grading practice and can explain it — but to point out how formalized work processes, particularly high(er)-stakes ones such as teaching observations, may create, without intentionality on the part of writing program or departmental administrators, resistance for instructors, especially contingent instructors. Even for GTAs, whose position is generally not at risk from a poor teaching observation, the inherent scrutiny involved in the teaching observation process, especially when coupled with perceived resistance or hostility on the part of the observer, may cause anxiety and interfere with free, agentive assessment choices going forward.

This example further illustrates how institutional processes may work against each other, at once providing support and resistance for ungrading among educators. UTK educators largely describe the program and in fact the wider institution as supportive of ungrading, but this does not change the fact that they find friction between their new habits and longstanding institutional processes such as observations. Power differentials make the restrictions on ungrading particularly apparent, as educators may wrestle with, on the one hand, apparent disapproval from a more highly-ranked (TT/T faculty, NTT faculty, or advanced GTA) observer and on the other encouragement and direct support from (a) peer(s) in the form of the Google folder.[77] In this

[76] In first-year writing, UTK uses NC to replace all non-passing grades (e.g. C-, Ds, F). See the catalog for more information: https://catalog.utk.edu/content.php?catoid=24&navoid=3078#ABC_NC_grading

[77] To be clear, graduate students and non-tenured instructors do not automatically *lack* confidence, as highly varying expectations around one's position and any reviews or observations or the absence of such could cause an instructor to have *more* confidence in ungrading. This is the case with some participants. This is also the case with me, as during the 2020-2021 academic year I was scheduled for a teaching observation and had the confidence, thanks to

case, GTAs or non-tenured instructors' relationship with me or with the Google folder makes ungrading visible as a potentially-valid option for assessment in FYC at UTK, only for longstanding architectural features of the program such as observations to raise questions about the validity of ungrading as an effective pedagogical choice. My point here is not to criticize observations, which especially for newer instructors such as GTAs, are often an important learning tool. Indeed, the fact that observations play a constructive role within the program is key, since it underscores that important, meaningful aspects of existing programmatic architecture, established without the intent to resist ungrading, may function in a way that proves hostile to innovative assessment practices[78]. The experience further highlights that during AY 2020-2021, UTK did not clearly support or depress ungrading, one way or the other. It did both, via systematized rhythms and hierarchies so embedded into the work processes of the department that they easily go unnoticed. The point, then, is quite simply that even institutions which aim to support educators, if they do not adequately reflect on and commit to renovating their preexisting architecture, may wind up undermining their own hoped-for support for ungrading.

Relationships

As the section above suggests, *relationships* are key to the institutional structure(s) that supported ungrading at UTK during AY 2020-2021. Though not explicitly discussed in institutional ethnography, relationships — the connections both horizontally among peers and vertically between educators and administrators, aligned with the institution — are nevertheless central to the method, given its interest in institutions as dynamic, "highly networked (La France p. 4) spaces. Supportive relationships open space within the architecture of the program for instructors to explore new assessment decisions, yet by the same token, more directive, hierarchical relationships in particular may close off that space, making the space hostile towards teachers' interest in taking up alternative, equitable forms of assessment.

The importance of strong relationships in affording space for ungrading is clear as soon as UTK educators start weighing their decision about which assessment method to use. Some educators expressed concern about whether "the department is on board" with a new assessment approach or whether they were going to "get in trouble" (Kerry). Such concerns underscore that in choosing to ungrade, or to stick with conventional grading, instructors look to their relationships. How will other people, some of whom are in hierarchical or authoritative positions, look at their use of ungrading? How will their relationships change? Reflecting Schwarz's finding that educators often associate permission to ungrade with support for the practice (2020), UTK instructors' valid concerns about their professional relationships drives — and in one case, restricts and alters — adoption of ungrading and perception of its use.

For the most part, peer and hierarchical relationships at the University of Tennessee, Knoxville supported contingent instructors' use of ungrading. Both Lane and Kerry point to their prior connection with me, a peer, as giving them confidence, stressing that I was "tangible [and] real"

---

my expertise not only in ungrading but more broadly in teaching, with more than a decade in the classroom, to redirect the relationship to advocate for ungrading.

[78] However, the UTK FYC program (while I was in the Assistant Director's role, during AY 2021-2022) relied on this data to develop a protocol to guide observations for instructors using ungrading, to ensure the processes was adequately valid and supportive of educators' agency

and “in their department” (Lane); they could “walk over and knock on [my] door” (Kerry) to ask questions about ungrading. The emphasis on proximity, both material and social or emotional, is key here, since instructors at other institutions who rely on scholarship do not (necessarily) have the benefit of access to an expert located in their context, and happy to answer questions about their context. Also important is a social connection. As a fellow graduate student in the same department, I am colleagues with the graduate students represented in this study; it is easier to knock on my door than on the door of a faculty member, who (regardless of how approachable they are) maintains a more hierarchical relationship of necessity. As fellow graduate students, our work also brings us into proximity with each other; we attend many of the same events, enroll in the same courses, and work on similar projects. Our peer relationship creates (at least, ideally) sufficient comfort and familiarity that Lane or other UTK instructors feel free to reach out with questions. This support in turn assures instructors that they will not “get in trouble” for using ungrading and gives them confidence in taking up and fruitfully deploying alternative assessments in their own classroom.

This does not mean, however, that relationships with faculty or programmatic and departmental directors do not matter. Far from it. Supportive relationships with individuals in more authoritative or hierarchical positions are also key to ensuring that the composition programs remain hospitable to the use of ungrading among contingent faculty. At UTK, at least one graduate student instructor spoke of their relationship with a faculty member as influential in supporting their use of alternative assessment methods. Jamie says:

> One of the reasons ungrading worked so well for me, particularly when it came to giving students feedback, was because I felt like I had a good example of it from one of my graduate instructors that I had a lot of classes with. But this instructor that I took multiple classes from had a very — maybe lax isn’t the right word — but a very flexible approach to grades, and I feel in particular the way that [the professor] responded to my work emphasized the relationship that the instructor and student had and engaging with the ideas and even engaging with Oh! that’s an interesting point. There’s nothing formative about that but you’re just having a conversation with the student’s work, right? and building rapport in that way, and treating them as an equal, and I feel like the example I had as a result of that enabled me to be a much more compassionate instructor who gave formative, conversational feedback to my students rather than punitive feedback.

A few things stand out about Jamie’s response. First is that Jamie took multiple classes from the professor in question, meaning that the professor is a person central to the student’s program, likely viewed (especially given the warmth of this description here) as a trusted mentor. The professor’s assessment choices carry with them a great deal of ethos; they are choices that the graduate student feels they can trust. Second is Jamie’s description of the professor’s approach to assessment as a “very flexible approach to grades,” a relational one which centered “engaging with the ideas” in a project over “punitive feedback.” Although the professor is not described as using ungrading per se, their overall model shares similar values, among them the desire to craft a relationship between teacher and student that centers mentorship and rhetorical development over penalizing students for perceived deficiencies. In this way, Jamie’s ’s experience with the professor transferred into their own teaching, opening space for ungrading as facilitating similar

ways of responding to students. This passage is important, since although many graduate faculty may not be familiar with the precise moves that ungrading makes, they can nevertheless support its use, or creativity and innovation in assessment among graduate instructors, simply through adopting non-hierarchical assessment practices in their own classrooms. The name(s) that we give to our assessment practices are, in the end, less important than the values or goals at which we aim; and having role models who share and exemplify what it looks like to aim at these goals in feedback and assessment is key.

Supportive relationships with programmatic directors as well as faculty is also key in ensuring the program remains hospitable to instructors' choice to use ungrading. Several participants described a collegial, trusting relationship with the First-Year Composition Director at UTK as giving them confidence to try ungrading. Kerry says notes that they feel confident in their use of ungrading knowing that the program director "thinks I'm doing okay." Kerry adds that it also gives them confidence to know that even "if something were to go weird and wrong, the head of the Comp Program would not be grouchy at me." Effectively, what Kerry describes here is trust. Kerry feels trusted by the Composition Director to exercise their own expertise and agency in the classroom without feeling a need to get things perfect the first time. From Kerry's perspective, the Director's trust in instructors, regardless of whether things go smoothly or "weird and wrong," is necessary for ungrading. Given the strong value UTK instructors place on agency in their teaching decisions, as the prior chapter indicated, having assurance that institutional leaders trust them to make effective pedagogical choices is empowering. The flip side of this, however, is that instructors who do not feel as supported by their writing program director(s) or know them well may feel less sure about their decision to ungrade. To be clear, not every instructor will "talk to [the program director] a lot," as Kerry describes; many instructors may not want to. Yet trust remains key in constructing a programmatic space hospitable to alternative forms of assessment. Knowing that the programmatic director or other institutional representative — including, I would suggest, higher-level admin — will not be grouchy about innovation or change is key in opening space for contingent faculty to do the messy work of making complex, creative changes to their assessment practices. Close down that space, express skepticism about contingent faculty's agency or surveil their choices in classroom teaching, and a program may become more hostile to ungrading.

In fact, at UTK, less supportive relationships constrain or alter instructors' decision to ungrade at constructing the program as more hostile within the context of those relationships. Speaking of their relationship with a faculty member serving in a supervisory role, a graduate student — as in the section on work processes, I am omitting the pseudonym for extra anonymity — describes receiving critical feedback on their planned approach, which shifted the assessment method away from ungrading towards more conventional approaches. As the graduate student instructor describes, their supervisor "doesn't really do contract grading [and] thinks that I need to weight some things, and put more weight on assignments and less weight on some other things." This direction in turn restricted the graduate student's flexibility in developing their assessment strategy:

> I didn't originally want to [weight assignments differently] but as a GTA, I don't feel like I have quite as much freedom to like choose the exact policy I want, because — and I like

> [supervisor's name] as an instructor, I'm not saying anything bad, but [they] want me to weight a couple of the categories so put more weight on assignments and I'm just not sure I want to do that, you know. But then it's like well, [they're] basically my mentor for [this course] so I probably will end up doing that, so there is I think a bit less freedom that I have with contract grading right now as a GTA, because you're still working under people and within the system.

I want to be clear here that at the time, alternative forms of assessment were a relatively new concept within the program. While a few instructors were familiar with them, many weren't. For faculty serving in advisory roles over graduate students, recommending caution about use of a new, largely untested (at least, within the UTK program) assessment method is not unwarranted. I am certainly not including this excerpt to criticize the supervisor discussed. What I do want to stress, however, is that the graduate student perceived the interaction as restricting their ability to ungrade and pushing them back towards a conventional grading approach, with more weight on assignments. The example illustrates how a *lack* of support may constrain instructors' use of ungrading, especially for graduate students and contingent faculty. Contingency causes instructors, whether graduate students or NTT faculty, to feel more acutely the pressure of "working under people and within the system" and making choices accordingly, shaping their agency to fit systemic expectations. At UTK, hierarchical relationships across the FYC program are largely diluted, as, though instructors teach from a standardized curriculum, they have the freedom to plan their own lessons and create their own rubrics — and, for the most part, create their own grading approach. Yet the hierarchical pressures do not go away and effectively draw in the landscape of possibility on some instructors, calling into question their use of ungrading in the class. Undermining instructors' desire for agency in their own teaching decisions, these relationships illustrate the power of connections across an institution to inhibit, as well as strengthen, the use of alternative assessment mechanisms. Programs interested in sustaining ungrading or similar approaches may want to consider options for advising and supporting teachers in ways that affirm their interest in alternative assessment practices.

Conclusion, 2020-2021

What I learned from focusing on UTK educators' experiences with ungrading through the lens of key concepts from institutional ethnography is that institutional architecture matters. Construct the program one way, with strong models for alternative assessment, work processes hospitable to instructors' decision to use ungrading, and educators lean into alternative assessment. Yet construct the programmatic architecture another way, one where architectural features — regardless of individual interest among faculty — are hostile to ungrading, and the use of alternative assessment may falter. The potential of hostile, or even neutral, programmatic architecture to undermine and discourage the use of ungrading is clear, given that, faced with pushback, instructors may choose not to ungrade. Overall as educators' responses make clear, UTK, even without any formal, departmentally-sponsored support in place for ungrading, largely welcomed the change. The features of its institutional architecture, from its ethos of non-authoritarian leadership to easy relationships among GTAs, which encouraged resource-sharing and pedagogical discussion, encouraged cultural change and allowed the grassroots movement to take root.

Yet institutional architecture is like an old house — carefully built, hard to change, and often resistant to new forms of assessment in surprising ways. Often, features of the programmatic architecture installed years previously to good intentions and results, such as yearly observations or mentoring relationships, provide a spacious environment — until faced with a new, innovative practice such as ungrading, which does not easily fit within the constructed space and encounters resistance. Adapting the term hostile architecture to include unintentional resistance or control exerted over a particular space, as well as intentional, I show in the pages above how, despite the encouragement, direct and indirect, that the Composition Program offered to instructors starting their ungrading journey, other aspects of the architecture generated resistance for instructors, slowing down or halting their process. Ordinary features of the programmatic architecture — among them mentoring relationships and teaching observations — that predate the movement towards ungrading may inadvertently construct the space in ways hostile to equitable assessment, restricting its growth. What this suggests, given the risk that contingent faculty or GTAs may face in taking up alternative assessment, is the importance of rebuilding the institutional architecture, through intentional, formal support more hospitable to educators' decision to use ungrading. I discuss the UTK First-Year Composition Program's efforts to do just this in the following section.

2021-2022 data

My next section revisits the question of how programmatic architecture may be hospitable towards the use of ungrading among contingent faculty, within the context of the 2021-2022 Equitable Assessment Pilot. I draw in this section on my analysis of the second round of interviews (2021-2022), with participants Finley, Jamie, Kerry, Oakley, Avery, Sterling, Taylor, River, and Greer. In approaching this data, I was encouraged by my initial research on institutional affordances, or infrastructure, supporting ungrading, and decide to conduct a more thorough study of the programmatic support for ungrading during AY 2021-2022, given that the second year was characterized by the more formal, or direct, support of the Equitable Assessment Pilot. Specifically, the Equitable Assessment Pilot supported instructors via a $200 stipend per term, along with regular check-ins and a start-of-term workshop to guide them in making decisions about their assessment practices. Because writing programs are large, complex, multi-tiered spaces, however, such direct support — an obvious indicator of administrators' interest in alternative assessment — is not automatically equivalent with a hospitable architecture. I wanted to know how direct support altered, if at all, the programmatic architecture in ways that made it more hospitable to instructors' willingness or confidence to use ungrading and their experiences with it. I also wanted to know how participants within the Equitable Assessment Pilot developed an assessment approach that aligned with their pedagogical values and what recommendations they had for further iterations of the Pilot, as the English Department and First-Year Composition Program intended to continue supporting the use of alternative assessment practices among writing faculty.

Accordingly, in analyzing the data, I paid attention to participants' responses about the writing program, collecting any mention of the writing program into a separate code for subsequent analysis. I then coded segments within the writing program code, looking for repeated words and themes and relying on key concepts of institutional ethnography, the inherently hierarchical, top-down nature of institutions contrasted with the inevitable bottom-up "social coordination"

(LaFrance 2019, pg. 31) on work activities and responsibilities among members of the institution, to guide my analysis. My analysis resulted in the following codes, which I include in Table 4.1, along with a brief description of the code's meaning or limits.

What I found in analyzing the interrelationship among these codes is that the permissory aspect of the pilot is key, as instances of resistance or friction with the institutional status quo do not appear, as they did in the 2020-2021 data. Ensured the freedom and safety to use ungrading, UTK educators took advantage of this hospitality, at once accessing both administrators, faculty, and other educators with ungrading expertise while also relying on peers and a mix of resources (inside and outside the program) to develop an approach to ungrading that aligned with their pedagogical identity and commitments. As I discuss, I include observations or recommendations that participants included for subsequent iterations of the Pilot. Some of these recommendations have already been adopted, as the Equitable Assessment Pilot has continued in AY 2022-2023 with support from the Student Success Grant at the University of Tennessee. Nevertheless, the recommendations are discussed here both as evidence of the development of a robust architectural support for alternative assessment practices, and instructor agency in developing those practices, at the University of Tennessee and in hopes that the recommendations may be useful to readers.

I tackle key themes in the section below, first discussing how participants perceive the Equitable Assessment Pilot as granting permission/freedom for ungrading and ensuring their safety in taking it up. I then turn to the impact of this permissory nature — a word drawn from participant interviews, capturing their clearly-articulated sense that the pilot granted permission for ungrading — on the development and experience of alternative assessment. In exploring this question, I show how the minimally hierarchical nature of the Equitable Assessment Pilot fosters a space hospitable to participants' use of administrative resources. Indeed, my second point is that the pilot is sufficiently hospitable to instructors' agency to choose and design their own ungrading approach that it includes the freedom to deviate from those resources, accessing other resources that cultivate their approach to assessment. Ultimately, it is this flexible, permissory nature which cultivates the UTK First-Year Composition Program as hospitable towards instructors' interest in ungrading, opening space for wider use of alternative assessment in first-year composition courses.

Permission

As I explained in the first half of this chapter, while 2020-2021 participants largely experience the UTK FYC program as hospitable for ungrading, a few architectural features generated resistance. The Equitable Assessment Pilot, however, turned programmatic support for alternative assessment from informal to formal or official, effectively renovating the institutional space.[79] Educators liked the change. The Pilot gave educators permission or freedom (also a

[79] I read much of the following analysis — instructors' sense of having permission to ungrade — as speaking of the psychological impact of legitimizing ungrading through a formal pilot, compared with simply "going rogue", as to some degree educators in the 2020-2021 data did. (Certainly I felt as though I was going rogue.) But it's important to note that because I had access to the 2020-2021 data as I planned for and contributed to the design of the Equitable Assessment Pilot, institutional work processes changed. I also revised and distributed a new set of teaching observation documents (for GTAs being observed by lecturers) to validate alternative assessment practices

Table 4.1 Codes With Definitions for Educator-Writing Program Interactions, 2021-2022

Code	Definition
Permission	Participants speak of the Equitable Assessment Pilot as granting "permission", "freedom", or "safety" to pursue alternative assessment practices.
Papers	Participants describe documents and resources they accessed in creating their ungrading / alternative assessment approach. Documents include models of ungrading approaches shared via a Google folder, academic scholarship, and invited talks. Participants may discuss a desire for access to more documents, as well as their use of existing documents.
Prior experiences	Participants discuss the effect of prior experiences with alternative forms of assessment on their use of ungrading. These prior experiences may include their own educational backgrounds or in their pre-Pilot use of alternative assessment practices/ungrading.
People	Participants discuss the people they spoke to or wanted to speak to in creating their alternative assessment model and using it over the course of the term. These people include experts on alternative assessment and related theories but also include peers and colleagues within the Pilot. Participants may also discuss how their use of ungrading and perception of the Pilot is altered by student experience.
Priorities and principles	Participants discuss how their goals and values as educators shaped their experience in and perception of the Pilot, along with the ungrading model they created.
Planning	Participants discuss how the Pilot contributed to or shaped their planning for an ungraded semester, along with describing features that would improve their planning in subsequent semesters.

and guide observers in providing constructive feedback. The 2020-2021 and 2021-2022 data both do and do not occur within the same institutional space, as the UTK program was already changing based on the research conducted in this study.

word used in participant interviews) to ungrade, encompassing both the freedom to choose without possible percussions to use alternative forms of assessment and the safety and support do so effectively. Avery, a graduate instructor, emphasized permission as a core affective experience of the pilot:

> It's such a funny thing to feel like unsure of the institutional reaction or response to different teaching methods and kind of what you're allowed to do. But I feel like — yeah, like I said, for me the pilot felt very permission-giving, like our meetings and discussions around it and being able to ask questions and hear what other people were doing was really helpful.

Key here is the contrast Avery draws between belonging to institutions more generally, and their experience in the pilot. Within an institution, as Avery points out, educators are likely to "feel unsure" about what pedagogical or assessment choices are permitted. This feeling is acute for graduate instructors, who may be only a year or two out of their bachelor's degree program and unused to making instructional decisions about their course. This institutional authority, however, is minimized within the pilot, which Avery describes as "very permission giving." Avery mentions the ongoing conversations, both collectively (meetings and discussions) and one-on-one (questions) as playing a role in the pilot's permissory character. Inviting participants to talk about ungrading, the pilot makes assessment less a matter of receiving institutional guidance, the boss texts of rubrics or grading scales created by the writing program or university and passed down, and more a matter of choice-making. Talking about assessment choices, legitimized them.

Greer, an NTT lecturer, echoes Avery's description of the pilot as permission giving, saying, "I probably always had permission to do that [use ungrading] but the pilot made it feel like I had permission to do that, so that was good." Greer's point that educators within the UTK First-Year Composition program "always had permission" to use ungrading is key. Indeed, this line stands out to me because I too worried about whether I had permission, before realizing (through a series of emails with the then-director) that so long as my assessment choices were transparent and fair, the Composition Office did not particularly care what practice I used. But implicit permission is not the same as explicit permission, and Greer's concern here picks up the reality that, in determining action within a hierarchical organization or institution, an individual's location matters (LaFrance, 2019). What is obviously permissible for somebody at the center of the organization is less obviously so for somebody at the outskirts. The pilot, then, is key, as it communicates the validity not only of alternative assessment generally but more specifically of instructors' agency to make choices about their assessment practices to members historically on the outskirts of the program, in this way flattening (at least a little) the institutional hierarchy around grades and grading choices. This freedom — to exercise their own pedagogical agency – – was well-received by instructors, as Greer notes about the pilot: "I'm an incredibly autonomous and independent person so anytime someone gives me even more space I'm like yes, awesome, leave me alone. So, I was pretty happy."

For instructors, "permission" meant not only authorization from the Composition Program to use equitable assessment but also professional security in doing so. Instructors could opt in to

ungrading knowing that, in making this shift, they would not be left to fend for themselves but would have resources, personnel, and colleagues to assist them with any challenges. River makes the importance of this security clear in their conversation, describing the reason for their decision to ungrade:

> It's not like this was the first I'd ever heard of it [ungrading], but it was the first time I'd ever thought like, yes, I'm going to do something more radical, I'm not just going to tweak what I'm already doing, I'm going to change what I'm doing and knowing that there were resources for me and that knowing that somebody who knew more about it than I did that I could ask questions to, I found that sort of safety net very useful …. Even if I didn't use [that safety net], knowing it was there gave me the push to do the thing, because I knew that if I was doing something new, and something started slipping sideways, I could course correct with someone who at least had more experience than I did.

What stands out to me about River's observation is the connection they make, between their affirmative decision to make the leap into ungrading and the "safety net" provided by the Equitable Assessment Pilot. Without that net, as River observes, instructors risk "something slipping sideways." The undefined "something" is important, as it gestures towards a sense that ungrading could go wrong in limitless and serious ways. For educators new to teaching and/or on the margins of power within an institution, both characteristics that apply to graduate student teachers, departing from norms is a scary step into the unknown. For River, the Equitable Assessment Pilot minimizes that fear and makes it possible to step into the unknown, by providing a net; the implication is that if an educator does in fact trip or fall while trying out ungrading, somebody with experience and authority will be there to catch them. Having somebody with experience is valuable for educators, as River points, noting that they wanted to be able to "ask questions to" somebody who knew how ungrading worked and had used it in the same courses, at the same institution. Other educators concur with River that "having [a] support system of other people" (Sterling) or "feeling like [they] could go to" experts with questions (Avery) is key in validating their decision to ungrade, though they lay less emphasis than River does in the decision-making role of that support, pushing them over the ledge into ungrading. In my discussion of the prior year of ungrading (2020-2021) in the UTK English Department, I called attention to the value of supportive relationships from both supervisors and peers in authorizing people to ungrade, and ensuring the program remained hospitable to their choices about alternative assessment. Those themes recur here, amplified, as the Equitable Assessment Pilot formalized and made explicit the support(s) that had previously been implicit. In this way, it's important to call attention both to the value of explicit, departmental authorization of ungrading and the safety net as formal changes to the programmatic scaffolding, effectively renovating the department into a place where instructors need to worry less about whether the choice to use equitable assessment is "allowed" and can focus on making assessment choices grounded in scholarship and student well-being.

One point I want to draw out is that for instructors, the permissory nature of the Pilot, its explicit validation of their decision to adopt assessment practices geared towards learning and equity, provides a powerful corrective to historic institutional norms. As my earlier chapter on the

historical roots of grading in the UTK First-Year Composition Program observed, the University often imposed, top-down, particular grade distributions, warning about the dangers of grade inflation. A consistent bell curve — not too many As, now! — puffed up its institutional prestige, the deliberate cultivation of reputation for rigor. To my knowledge, neither the UTK First-Year Composition Program nor the wider institution collects grade distributions on a regular basis, though of course such information is available to select administrators, and similar information may be collected as part of regular evaluations. (For instance, the GTA Teaching Evaluation in the UTK First-Year Composition Program asks instructors to submit for their evaluator's review four anonymized graded papers at a perceived A-level, a B-level, a C-level, and an F/NC level.) The assumption that instructors are, of course, maintaining the bell curve and of its status as a marker of accurate grading persists, even a hundred years later. As Taylor points out, the Pilot provides an important buffer both against this assumption and against any possible institutional ramifications from the use of an alternative assessment practice. Speaking of the Pilot, Taylor says:

> I do think that you would have to have active support from your writing program to be able to do this [adopt ungrading] in any kind of way that felt good or that felt supported, like your department was going to stand by that — I mean, even just looking at my grade distribution, if I wasn't teaching in a program where that was — where people knew that this pilot was happening and that I was using equitable assessment and ungrading — I would feel like less confident about that grade distribution, because I would worry, maybe especially as a non-tenure track faculty member that there would be some kind of repercussion or somebody would say something about how many As I was giving.

What Taylor's observation gets at is the reality that, thanks to the institutional histories around bell curves as normative, such a distribution is commonplace as a marker of quality education. Instructors who deviate from that norm, even if they deviate by giving out *more* As, may have their ability as educators called into question. As Taylor alludes to, they (like many of their colleagues) found that using ungrading raised their grade distributions, sometimes significantly; educators tended to give out more As using ungrading, than using conventional methods of assessment. Yet the commonplace of a bell curve is sufficiently normal that, especially for educators like Taylor or Greer who inhabit a position towards the edges of the institutional landscape, rather than the center, they may experience pushback, skepticism, or even repercussions for such an outcome. Think for instance of the GTA from earlier in this chapter, who, though they were not to my knowledge challenged about the grade distributions in their course, did receive pushback from their observer on the decision to use ungrading in the first place; an A-heavy distribution could add weight to potential criticisms.

In this way, the Pilot provides an important role in granting institutional authority to ungrading as a legitimate decision, blunting potential criticisms, rooted in its own institutional decisions of maintaining rigor over instructor and student well-being. Put another way, though historical institutional traditions of prioritizing rigor over equity in grading cannot be undone, the Pilot (re)structures or renovates the institutional space in ways that close off, as Taylor notes, potential critiques of instructors based on historical norms. By authorizing ungrading as a legitimate decision, the pilot instead makes the Composition Program hospitable to contingent instructors

looking to center equity in assessment — which may, in turn, alter the grading distribution, rooted as the bell curve is in *in*equity. This decision is what allows for, ultimately, the bottom-up, grassroots turn towards ungrading that the UTK FYC program experienced during 2021-2022.

Limitations to the permissory nature of the Equitable Assessment Pilot

Despite the fact that instructors experienced the Equitable Assessment Pilot as granting both the permission and safety to ungrade, and its role in (re)shaping institutional architectures to prioritize instructors' decision to center equity in classroom assessment, the hierarchical nature of belonging to and grading students within the First-Year Composition Program did not go away in instructors' experience. The ecological, living nature of institutions, their continual (re)constitution through dynamic, interlocking networks, makes hierarchy difficult (impossible) to abolish altogether. One form of this hierarchy is the existing FYC curriculum, which instructors found interfered with their use of more liberatory assessment practices. Finley calls attention to for instance the emphasis on "citation styles in the way that ENGL 101 is structured," noting: "I can have a conversation about why citation is important but why do [students] need to know exactly how to do APA and then in two weeks know exactly how to do MLA?" Importantly, the ENGL 101 and ENGL 102 curricula have existed in their current form since 2008 and predate both the current and previous Directors of Composition, meaning that they represent not a particular individual with authority but the larger, more enduring and complex, hierarchical authority of the institution.

This authority, nevertheless, proves hostile to UTK instructors' desire to move towards equity. Reflecting on how the emphasis on citation can box students into ways of writing oriented around following clear, receive guidelines, Finley concludes that it's "the kind of thing where I'm like, all right, well, I'm doing this ungrading and that's cool, but [ungrading is] crossing with the curriculum." I read Finley here as expressing frustration with the friction between their interest in following the values set out by ungrading — among them, exchanging a carceral, "punitive" approach for one that emphasizes more creative, "flexibl[e]" approaches for students — and the continued emphasis on correctness at some places in the curriculum. As a contingent educator, however, Finely likely does not feel free to change the curriculum, which they receive from the Composition Office via syllabus or Canvas templates. The curriculum likely has, for Finley as for other contingent educators, an aura of authority — something they are required to use, as a condition of their job.[80] (Nor, for that matter, are contingent educators that interested in altering the curriculum, a time-consuming, uncompensated endeavor.) Curriculum, of course,

[80] I suspect, based on conversations with my advisor during revisions, that writing program administrators and/or long-term NTT faculty or TT/T faculty may perceive curricula as potentially hostile or resistant to ungrading not because the curriculum has authority but because the curriculum has longevity. In other words, because curricula are hard to change, even at institutions with minimal bureaucracy, writing teachers and administrators may take up ungrading, only to find tensions between its values and the values articulated in the curriculum. They may further find the curriculum resistant to change given its entanglement with other features of the programmatic architecture. I also suspect, however, that GTAs and more recently-arrived NTT faculty may nevertheless perceive the curriculum as authoritative, a difference I attribute to a feature key to IE, the importance of standpoint in interpreting institutional activity (LaFrance, 2019). Contingent educators — GTAs in particular, who are by definition temporary workers — are likely less invested in the history of the curriculum within the institution and more likely to experience it as simply a natural feature of the landscape, something they have to do but which nevertheless creates hostility or resistance towards their use of ungrading.

changes slowly, especially at a large, public university, and consistency in curricular choices from one section to another is key in ensuring students have a positive, equitable educational experience across the institution. Yet the point remains that despite explicit institutional support for ungrading and attention to renovating institutional values, the common course objectives educators feel required to teach continue to shape their decisions about what to spend time on in class and may limit their ability to lean into the potential goals and directions suggested by their new use of ungrading. This restriction underscores the need for the University of Tennessee to continue renovating its architecture to be more hospitable towards ungrading, reflecting on and reworking its values across the first-year writing curriculum.

A second, and perhaps more intractable, form of institutional hierarchy that affects educators' use of ungrading is that the First-Year Composition program (like all composition programs) exists within other, larger hierarchies, in particular the University of Tennessee. Whatever changes have occurred within the First-Year Composition Program, the University remains committed to conventional grades, in turn cultivating a wider, hostile architecture for the use of ungrading among first-year composition instructors. This commitment is unsurprising, since as I observe in the historical chapter, grades serve the institution, providing a shorthand for transmitting information about student performance within and outside of the institution and imposing its own value system within its borders (Scott, 1999). But the intractable nature of grades within the larger institution — their ongoing systemic role, and the reality they are likely to be much more difficult, if not impossible, to alter or eliminate — does not change the fact that instructors find them an immediately present barrier to the goals of the Equitable Assessment Pilot. Reflecting on drawbacks of the Pilot, Avery says:

> We have to work in a system that still ultimately grades everyone, and I really struggle with that, like translating like labor-based grading into any kind of number at the end. [Students are] still having — coming out of this with oh, I get X or Y grade versus like a straight up kind of pass / fail …. So yeah, I think — I don't know, obviously that would be a much larger university policy shift to move away from that entirely, but it's just so clear to me the benefits of this, it felt really true in my class, so then when it came to the end of the semester to be like, *And now here's your grade!* you know? And we did that activity in the end of the semester where each student was kind of working out what their grade would probably be, so I don't think anyone was surprised by their grade or anything like that, but I think it was still uncomfortable to have that sense of an evaluation after a whole semester of not having that, so it was a shift for all of us, and you know, I yeah, I think at this point without like a larger shift that's not avoidable.

Avery is among the instructors who spoke explicitly of the value the Equitable Assessment Pilot had in authorizing instructors to use ungrading, and here, they discuss ungrading in warm terms, emphasizing its fit and appropriateness for their class. It's clear that for Avery, the Pilot opens space to put students and student learning first. At the same time, however, Avery situates that appreciation and greater feeling of freedom within the First-Year Composition Program, within the larger system of the University itself, which "ultimately grades everyone" and requires that instructors "translat[e] labor-based grading into a kind of number at the end." As Avery describes it, that final evaluation, especially after a semester of ungrading, is a cold shock, like

jumping into a pool; and the discomfort and ill-fitting connection, between more liberatory forms of assessment and the conventional letter grades required for institutional records, undermines the goals of the Equitable Assessment Pilot, and the teachers working within it. I don't want to be too grim here. After all, Avery turns from discussing the limitations imposed by the larger University grading requirements to expressing appreciation for the value of the Pilot in making possible more liberatory forms of assessment, grounded in instructors' own agency and ability to choose wisely for their classroom. And certainly the University of Tennessee is not unique in requiring educators to submit final grades for their students. Only a very few universities — often small, private colleges such as Reed (Blum 2020) — do not ask for final letter grades. But as I turn to discuss the resources and materials that instructors relied on to pull their ungrading practice together, showing how the minimal hierarchy cultivated by the Equitable Assessment Pilot made possible the kind of wide-reaching, collaborative work key to a grassroots movement, I also want to be realistic that the hierarchy is not gone, only offset and reworked in ways less hostile or threatening. It persists to shape and direct instructors' choices, even in ways that, as Avery points out, are hostile towards instructors' choice to use more liberatory, antiracist forms of assessment.

Instructor Use of Administrative Resources

Because the Equitable Assessment Pilot (re)shaped the program into a minimally hierarchical space, becoming hospitable to (ideally) student-centered choices, instructors exchanged a more passive receipt of assessment materials from the program for a more active, participatory engagement with these materials. For expert advice, instructors looked to the administrative center — in other words, the Composition Office, the administrator(s) who worked there, and the materials provided, from syllabus templates to ungrading models designed for the local context. To personalize and adapt their ungrading approaches in line with their values and pedagogical commitments, instructors often turned outside the administrative center, drawing on colleagues, peers, and academic or practical resources, including those from outside the institution. A more passive approach to assessment is common in writing programs generally. As I observe in my earlier discussion of boss texts, writing programs often choose to make things easy for instructors, providing material such as syllabus templates, Canvas shells, assignment directions, and (sometimes) rubrics for instructors to use in their own classes, often without adaptation. Indeed, in providing these materials to instructors, UTK is not even particularly rule-governed. In contrast with programs (Dively 2015) which go so far as to provide a menu of lesson plans and activities for instructors to use, UTK leaves instructors a lot of latitude around the lessons they develop, the grading rubrics (ungraded or otherwise) they use, and the themes they pick. as other programs. My goal in bringing this up is not to criticize the choice by some writing programs to more closely restrict their teachers' options in planning lessons and grading assignments; programmatic decisions are best made by those who inhabit the local context. My goal is to stress, however, that many writing programs — even UTK, to some degree — are run like a conical tent, with a single central pole; as the tent fabric takes its shape, flowing downward off the pole, so the structure and activity within writing programs originates, often, from the administrative center.. But, thanks to the permissory nature of the Equitable Assessment Pilot, that hierarchical, top-down nature is backgrounded in the University of Tennessee First-Year Composition Program, at least in terms of instructors' assessment choices. The result is not only that instructors feel free and safe to use ungrading, but that they also feel free to rely on materials

and people outside the administrative center, crafting their assessment practice in a way aligned with their goals.

This is not to say that instructors ignore the Composition Program altogether. They do not. Instructors made use of institutional or semi-institutional resources (such as my own ungrading folder) in two cases: when they wanted expert input, particularly for theoretically complex subjects; and when they wanted tried-and-tested information or material. The ungrading folder, which I described as a kind of quasi boss text, in that it supplemented and stood in for syllabus language about assessment, again proved helpful to participants during the pilot. Importantly, as the Equitable Assessment Pilot went on, I added (with permission) examples of strong ungrading models from other instructors who participated in the pilot, so that the folder represented a range of possible, tested ungrading resources for use in first-year composition courses[81]. Instructors widely spoke of the value of these resources. River, describing the materials they used in developing their ungrading practice, says: "In order to overcome the reservation of eek, I don't know what I'm doing, I relied on some of the materials that were already there, just to see how it went." They add, "I made some minor tweaks, but I didn't do a ton of innovating on my own, I just wanted to go with a model that I understood." River is certainly not alone in this decision. Oakley also describes "us[ing] the materials that you had already created, cause I figured that you knew a lot better than me what you were doing" and Avery describes the materials collected in the ungrading folder as a key "jumping off point to imagining" how to make alternative assessment work. What stands out to me about these responses is two things. First, instructors appreciate that they understand and know how the resources work. Especially for instructors who may be interested in alternative assessment but struggling to "imagine" what it could look like, having a concrete set of documents is invaluable in showing what ungrading actually turns out looking like. As Avery suggests, once instructors have that clarity, they're in a position to make informed decisions, "jumping off" into variations and developments that jive with their own values and goals. Second, as Oakley observes, the resources collected in the folder represent, especially for beginners, an expert point of view. Because I, and the other instructors whose material was included in the folder, have experience based on one or more semesters of using ungrading, beginners who feel uncertain or overwhelmed by possible choices are likely to lean on the resources we provide. As I discuss earlier, the ability to depend on expert materials is key to providing the security and reinforcement that makes the choice to ungrade possible in the first place.

The reality leads me to my second point, as, in addition to relying on tried-and-tested materials to develop their own ungrading resources, instructors also relied on institutional figures — in our case, faculty with particular relevant expertise, though staff with similar expertise, administrators, and peer colleagues could also serve — for guidance on theoretically or practically complex subjects. The key example here is the relevance of linguistic diversity and justice to antiracist assessment. As I gained experience talking about ungrading, and improved at facilitating these conversations, I grew careful about distinguishing ungrading as a practice from antiracist assessment. Antiracist assessment demands a working knowledge of sociolinguistics –

[81] I am not sharing other instructors' materials here, as the participation in the pilot was distinct from participation in the research, and I do not have, nor did I seek, permission to use their materials for this research study.

– in particular, the interconnection of language with identity and cultural privileging of white, standardized forms of English over other Englishes, such as African American Vernacular (Gere et al, 2021). While writing educators at UTK knew that ungrading has the *potential* to be antiracist, likely thanks to my own (early) emphasis on the connection between alternative assessment and equity, the actual theories and practice of antiracist assessment proved more difficult to access. First-year writing educators looking to develop their ungrading practice as antiracist leaned on resources and experts within the program for advice, especially a linguistics professor who, through AY 2021-2022, advocated for a more robust concept of linguistic diversity in the department. Taylor recalls the importance of these resources to supporting their assessment choices:

> With the linguistics stuff — and I know that although maybe [Professor's Name] is leaving we'll have visibly less [of linguistics stuff] at UT, but I know that UT has been thinking about this a lot, and you have also talked about this in your work, as well. Yeah. Cause that is like the only thing about the ungrading conversation that I really feel that I can't do without a lot of training from an expert about how to talk about like native Englishes as — being in the position that I am in, so that's the only thing that I felt like I couldn't really do without a lot more training or expertise.

As Taylor discusses here, they know enough about sociolinguistics to grasp the basic concepts such as multiple Englishes and the value of linguistic diversity. But their lack of expertise in the field means that talking with students about these concepts, or grounding their assessment practice in them, is more difficult, needing expert input. While the professor Taylor references subsequently left for another opportunity, their presence in the department, and the reliance on their expertise, underscores that one reason educators turn towards the administrative center is to benefit from more specialized knowledge on aspects of or values associated with ungrading and alternative assessment practices[82]. Though the ungrading folder, that quasi institutional document, makes changing assessment practices relatively straightforward, assessment remains a complex subject, which if instructors want to develop and improve, for instance by ensuring it works for racial equity, as Taylor discusses here, requires outside, expert input.

Instructor Departure from Administrative Resources and Guidance

Yet this turn towards the administrative center is also accompanied in the UTK Composition Program by a corresponding turn *away* from the administrative center, towards resources unconnected with the programmatic hierarchy. This is not a rejection of administrative guidance, especially in light of participants' use of core documents such as the ungrading folder. What it does suggest, however, is that the minimally hierarchical nature of the pilot made it possible (or at least desirable) for instructors to engage with other resources perceived equally useful to construct assessment practices that reflected their own values. To be clear, because the pilot was in its infancy, participants were unable to make as full use of these resources as they were of the

[82] Notably, Finley also spoke of the need for additional instruction in linguistics, suggesting an asynchronous online course that could be completed by individual instructors. That these theme recurs among two instructors suggests that the antiracist aspects of ungrading — its openness to other ways of languaging besides standardized American English — may be among the more difficult concepts for new instructors to grasp.

administrative resources available to them, quite simply because the structures were not in place to facilitate the ready access to non-institutional resources. But the insight is important nonetheless, as it suggests and moves towards a vision of equitable assessment in which educators have more autonomy in designing assessment practices that work for their pedagogical commitments and style. Where mentoring or teaching observations may prove (inadvertently) hostile to ungrading, especially if oriented towards grading as the norm for effective teaching, autonomy, or the ability to look to resources and professional networks outside the administrative center, is perceived among instructors as desirable, and when cultivated, contributes to a more hospitable programmatic architecture. I trace in the following paragraph a "wish list" that UTK instructors had for engaging with external or non-administrative resources, to show the importance of this autonomy to building architectural spaces that invite instructors' decision to ungrade.

River's experience is key here. River both expresses appreciation for the resources made available through the pilot and a desire for more detailed resources outside of those provided by or associated with the Writing Program. Having more scholarship on ungrading, River notes, would be useful:

> So maybe having more of a bibliography, if you're interested in antiracist equitable assessment here are the things to read, if you're interested in this kind, to help me get a better grasp of what's under this umbrella, what are the different methodologies, what — getting a sense of the field a little more because I'm new and wading in and there's so much to read, so I — I would personally find a sort of curated bibliography really useful so that depending on what I wanted to know I could better self-service a little bit.

River also points to the value of talking with other educators in the Equitable Assessment practice about their choices. In fact, River points to these conversations as sparking their interest in ungrading in the future, noting that they were "hearing from people who joined the pilot earlier than I did, so a lot of it was the environment here at UTK of having the pilot and also having other things outside the pilot, it was in the water, and so I was exposed to the ideas about ungrading." River similarly expresses a desire to continue these conversations, or continue learning from other people, in accessing their resources in the ungrading materials provided to people in the pilot:

> There are lots of different models that are sometimes quite different and I want to know what are some of the -- what are some of the things that would make somebody choose, Model A, B, C? Or again, what are the principles that underlie these things? Or, if I want to use equitable assessment specifically to do things that are antiracist for example?

Notice especially that in speaking both of the value of learning from outside research, such as in a curated bibliography, and in learning from other people's choices about ungrading, River highlights a desire to ensure their assessment choices are equitable and antiracist. I highlight this because River's interest in ensuring their choices are indeed equitable, and their corresponding desire to make use of resources outside the administrative center to guide their choices, links up with the argument I make in the next chapter, that educators' use of ungrading is grounded in

their values and often honed over a period of multiple semesters. The importance of time to ungrading — repeated use of, and experimentation with, a practice semester-by-semester — means that it makes sense as River suggests here, to ensure that additional materials, whether scholarship or peers, are available for educators to access, grapple with, and feed back into their own practices. Instructors who make use of non-administrative resources, as River wants to, are able to "self-service," selecting and rejecting from among a menu of texts that correspond with their own teaching needs and values, and help them think more critically about how to reach equitable or antiracist goals in their own practice.

Other instructors, including those with substantial experience of their own, concur about the importance of talking with other educators. Sterling, for instance, says this:

> I would like to know a little bit about what are other people's experiences, and what types of materials did they use for development, and how did that go for them? To maybe think about what I could gain from other people's use of ungrading, just to have more experience to draw on…. So for example something that I did in the first week of the semester was try to source some readings that were about concepts like language ideology and literacy and standard academic American English to have more of a conceptual language to give students to be able to talk about assessment. And I would love to know, like how other instructors maybe approached that, or to be able to share what I found helpful.

What is striking about Sterling's experience is that at this time, Sterling has multiple years' experience ungrading. That they still hope to listen to and "gain [something] from" learning about their peers' use of ungrading and choices in ungrading points to the importance of these horizontal relationships in sustaining and developing equitable relationships among first-year composition educators. At the same time, Sterling also speaks of a desire to "share what I found helpful" on equity and antiracism in assessment practices. Especially in light of River's shared desire to talk with peers and hear the rationale for their ungrading choices, Sterling's commentary here underscores the value of social coordination along the margins or edges of the institutional landscape. To involve teachers with antiracist assessment practices, help them not only *do* the right thing but know *why* they are doing it and align it with their own pedagogical commitments and values, requires this kind of peer-to-peer connection fostered in the UTK Equitable Assessment Pilot.

I believe overall that the minimally hierarchical, hospitable model of the UTK Pilot is a valuable one, as it simultaneously provides educators the needed resources to kickstart their practice and invites them to coordinate with peers and outside materials to self-direct their practice towards more equitable outcomes, with minimal regard to the constraints of the University. But I do want to be clear about one potential drawback. UTK is a predominantly white university (PWI), and the English Department has struggled to recruit and retain graduate students and faculty of color. Several instructors in the pilot expressed uncertainty about whether and how their ungrading practice could be antiracist in these contexts, or how to know whether their ungrading practice was antiracist, and though I did not speak to anybody who dismissed or undermined the importance of antiracism in writing assessment, it is entirely possible that some educators may

not see the need for this work, or even resist it. Additionally, scholars and teachers have identified Black faculty, or faculty from other minoritized groups, as facing the most risk in taking up alternative forms of assessment (Craig, 2021, McCloud, 2023). Going forward, I suggest that it is crucial for both UTK and other writing programs encouraging pedagogical innovation among teaching faculty as a way to sustain movement towards more equitable assessment practices to ensure that both administrative and non-administrative resources stress the importance of explicit, direct antiracist action and outline steps to make this happen in a way that invites choice and minimizes risks for all contingent or vulnerable faculty.

Conclusion

What I show in this chapter is that programmatic architecture — whether it is hostile or hospitable, or a mix of both — plays a key role in whether alternative forms of assessment take root in a program and grow, especially among contingent faculty. My point is not that ungrading is doomed if a writing program is arranged in a way hostile towards educators' agency in choosing alternative means of assessment, nothing quite so strong. But as I think of the GTAs and NTT faculty I talked to over the course of this study — Kerry and Taylor, River and Avery, and others — I think of the way they consistently speak to the important role the writing program plays in their experiences ungrading. Many instructors have experiences that invite them further into ungrading — the security in affirming relationships, especially with supervisors; the power of strong models, including (increasingly) from peers and colleagues as well as the Composition Program; ready access to specific expertise, including on antiracist elements of the program. Even the existence of the Pilot plays a role, given that for NTT faculty especially, it legitimizes grade distributions that deviate from the standardized bell curve towards As and Bs. Would the GTAs and NTT faculty in this study have used ungrading if the UTK First-Year Composition Program were not hospitable towards instructors' agency, to choose more risky, liberatory forms of assessment? Perhaps. I feel confident, however, in saying that the hospitality of the program towards ungrading ensured instructors used it in more creative, expansive, and interested ways, and that this same hospitality contributed to (hopefully) the long-term persistence of alternative forms of assessment within the Composition Program. Free to engage with resources on ungrading and experiment in their classes, instructors took up alternative forms of assessment more readily, with less thought to possible repercussions on their own locatedness within the wider program or professional and personal well-being.

The reverse, of course, is also true. Especially during the first year, when the First-Year Composition Program and the English Department at UTK provided only informal support, instructors found the program more hostile, or resistant, towards their agency in choosing ungrading. Even features of the programmatic architecture that had nothing to do with ungrading — yearly observations or mentoring relationships — interfered with instructors' sense of agency and security in taking up risky, alternative forms of assessment. This is not, of course, to say that these features of the programmatic architecture are *bad*. Indeed, as I took pains to point out in this chapter, programmatic architecture may be hostile to alternative forms of assessment in ways that exceed or bypass intentionality; institutions, and the people who work within them, often do not *intend* hostility towards alternative forms of assessment. This point is key, since it pushes writing educators, administrators, and other university leaders to consider the broader structural constraints that may be placed on educators and lower administrators looking to take up

liberatory forms of assessment. An interest in alternative forms of assessment is not enough. Nor is easy access to examples. The more committed leadership is to antiracist forms of assessment, the more leadership is obligated to (re)consider how their program may be renovated and reshaped to invite in alternative forms of assessment.

Looking back, I am astonished and grateful for the support that the University of Tennessee, Knoxville First-Year Composition Program provided, first to my own interest in ungrading and then to the swiftly-expanding initiatives to support ungrading across the program. The UTK FYC program had long tacitly allowed instructors to use alternative forms of assessment, so long as they were clear and fair for students; the Pilot made this tacit approval direct, and ensured instructors were able to do a good job. This architectural change in turn made space for instructors to both access programmatic resources, among them expert resources and faculty, and also to exceed them, drawing on non-institutional resources (or at least expressing a desire for such resources) to construct workable assessment designs that fit their own pedagogical values. Whether the formal structure of the Equitable Assessment Pilot, in contrast with the informal support of 2020-2021, caused a marked difference is not clear. The study was not designed to test the impact of the Equitable Assessment Pilot; it was not even originally designed as a two-part study. What is clear, however, is that the UTK First-Year Composition Program, especially (though not only) during the Equitable Assessment Pilot, renovated the institutional architecture in such a way that instructors felt authorized to choose ungrading, and safe and supported in their use of it. Many features of the preexisting institutional landscape or architecture contributed to programmatic hospitality — among them, as I have argued, an openness on the part of administrators to valuing instructor agency and experimentation, strong peer and hierarchical relationships, and ready access to resources. Other features, such as the protection and permission granted by the Pilot, were new. Together, however, they cultivated an institutional architecture open to instructor agency and decision-making, where they felt free to follow through on their teacherly commitments, take up alternative forms of assessment, and see those forms of assessment bear fruit. In the next chapter, then, I turn to the results or outcomes of a hospitable programmatic architecture, tracing individual teachers' experiences to demonstrate how the outcomes of ungrading are closely linked with teachers' values and realized more fully, the longer instructors use ungrading in their courses.

Chapter 5: The Impact of Teachers' Values on Perceived Affordances of Ungrading

So far in this project, I have focused on the institutional contexts for writing assessment — the role of institutional mission and reputation in imposing forms of assessment that prioritize standardized, white forms of language over students' identity and learning; and the need for a programmatic architecture hospitable to instructors' agency in facilitating the growth of ungrading. I turn now to discuss individual teachers' experiences with ungrading, drawing on my analysis of focus groups and interviews with contingent faculty at UTK to show how ungrading is rooted in the values and beliefs that educators bring into the classroom. In beginning this project with attention to the institutional levels of ungrading, I suggest those levels are more important to sustaining liberatory forms of assessment; it is not possible (or at least, quite difficult) to maintain long-term interest in ungrading in a hostile environment. Yet in turning now to the more individual level, I explore how contingent educators, working within a largely hospitable environment, experience ungrading — what outcomes they see from it, why those outcomes vary teacher-to-teacher, and what happens when they use ungrading over a longer period of time. This work clarifies how we as writing teachers (often our first professional identity, prior to and deeper than administrators or researchers) might take advantage of welcoming programmatic spaces to cultivate fruitful, liberatory forms of assessment.

As I discussed in the Introduction (many pages ago now!), research on ungrading — specifically on contract grading, by far the most common approach in rhetoric and composition — typically makes two, related assumptions about the effects of the model: that its primary effects are on students, and that it is primarily ungrading which exerts this impact. First, much of the literature on ungrading focuses on what the practice can do for *students*. Scholars conclude that contract grading have concluded that research improves student learning, making it possible (among other outcomes) for them to take risks and engage more fully with the writing-and-revision process (Danielewicz and Elbow, 2008, Klotz and Whithaus, 2021), that it gives them more agency over what they write about and how they write about it (Mallette and Hawks, 2020), that it gives them agency over how they participate (Gomes et al, 2020), that it encourages them to try out other ways of languaging more aligned with their identity (Inoue, 2019, Klotz and Whithaus, 2021), and that it corrects for the impact of racial bias, effectively restructuring the classroom to wall off or blunt the impact of teachers' linguistic biases, so that students are not penalized for a perceived failure to align with white norms (Inoue 2015, p. 128). This is important, crucial research, expanding our understanding as writing teachers and administrators of how assessment may be used to cultivate alternative, more effective, just, and hopeful ways of being, in the writing classroom. Yet in each case, the question asked and answered focuses primarily on students as the beneficiaries of alternative assessment: What do students get out of this? This list is not exhaustive, of course. Some work, especially the reflective pieces that emerged in the *Journal of Writing Assessment* special issue (13.2, 2020), argue that using, or adapting, contract grading and related forms of alternative assessment sparked reflection and reassessment for teachers, for instance inviting them to rethink their feedback practices (Wood, 2020) or grapple with their own raciolinguistic identity as teachers and their relationship to grades and grading (Craig 2020, Tinoco et al 2020). The dominant theme, however, is interest in contract grading as benefiting students.

Second and relatedly, the research focuses on what *ungrading* can do for students, framing it as the locus of change within the classroom assessment ecology. As in the paragraph above, I want to thread the needle carefully here. No scholar argues that ungrading alone causes certain outcomes, or that ungrading always causes certain outcomes, or even that ungrading *mostly* causes certain outcomes; the research is clear that like any other form of assessment, ungrading is highly local, contingent, and subject to varied outcomes (see Inoue, 2016, Litterio, 2016, Litterio 2018). Still, while ungrading *is* in many ways an intervention, a new practice educators take up in hopes of some change, emphasis on the impact of ungrading, or ungrading as the locus of change, has left largely unaddressed an equally important question: how alternative assessment not only affects the classroom experience but is itself affected, by the values and experiences teachers bring to the work. Accounting more fully for ungrading as a malleable practice is crucial, as it can help us grapple with the reasons for ungrading's success in our teaching — and also for its failure. Not all educators find that ungrading has the outcomes it is purported to have, as I noted in my Introduction, citing a teacher who speaks wistfully of using Fs to scare students into submitting (presumptively) stronger work (Inman and Powell, 2018, p. 46). I don't want to overstate this case, especially as the article overall finds mixed results with ungrading — some positive learning-centered outcomes, despite difficulty switching into a new system or way of doing assessment, among students and teachers both. But I do think that the statement points towards the ways that even when educators adopt methods of assessment that are, on their face, better for teaching and for justice, the educators themselves may remain committed to more punitive notions of their own work, judging students against an arbitrary standard. I would like to suggest that such commitments, though Inman and Powell do not discuss them, could account in part for the mixed results of the study. After all, if instructors still feel like cops, then is it any wonder students continue to feel anxious about their position and identity within that system?

In this chapter I take seriously ungrading's role as part of the larger classroom architecture. If ungrading, reaches out and (re)shapes that system, then ungrading must also be open to (re)shaping by the other components of the ecology, including — as I discuss here — instructors themselves, and the experiences and values they bring to the work of ungrading. Assessment, we know, is a deeply value-laden exercise, a working out and affirming of what teachers, administrators, and the public believe is important about writing (Broad, Huot). What I suggest is that though ungrading does have valuable affordances in the composition classroom, those affordances vary, based on instructors' own values and commitments in the classroom. Put another way, ungrading's outcomes are ecologically dependent. This work builds on research framing contract grading as a genre of assessments, meaning that educators take up and rearrange its various components based on their own values and beliefs about writing; contract grading is not one thing and should not be framed that way (Schwarz 2020). By exploring what values, in particular, motivate instructor uptake of ungrading and how those values translate into divergent experiences with ungrading and perspectives on the affordances it offers, I provide researchers and teachers a clearer sense of the mechanism by which ungrading (re)shapes the classroom and invite ongoing reflection and commitment on the values we bring to the work.

As previous chapters have shown, grades and grading in first-year composition have historically been a top-down practice, imposed through standardized syllabuses and aligned with university

mandates. In this way, grades and grading asserted university values — rigor and elitism, often at the expense of learning or student diversity. In the 1920s and 1930s, at the program's origin, the university claimed a mission of public service, yet the use of grades in first-year composition, including penalties for perceived grammatical errors, reoriented the composition curriculum around students' production of a single, platonic ideal of "good writing," the then-emerging standardized American English. Grades ranked students' performance against that ideal, relying on (artificially-maintained) bell curves to frame the wider institution as academically rigorous. First-year composition served a gatekeeping function, marginalizing students whose work did not align with the expected ideal. This function, though minimized by curricular and assessment revisions over the following years, remained sticky, as any top-down form of grading — even more minimal forms, such as syllabus templates — imposes programmatic or institutional values.

I also have argued that at the programmatic or institutional level, a hospitable architecture is necessary to cultivate a long-term, sustainable movement towards ungrading. Too often, writing programs are (inadvertently) hostile towards ungrading, as old nooks — such as teaching observations or mentoring relationships — create unexpected resistance for contingent faculty interested in taking up new, more liberatory forms of assessment. Yet by reflecting on where those points of resistance may occur, and attending carefully to the relationships, work processes, and boss texts or resources that the program makes available to educators, administrators can develop an environment more hospitable towards ungrading, laying the groundwork for a grassroots assessment practice..

This chapter shifts from the institutional to the individual level of analysis. I argue that within a programmatic architecture largely hospitable to ungrading, the practice may be fluid, not static, its outcomes dependent on the values and beliefs about writing and writing assessment that teachers bring to the work. Beginning with a brief overview of how COVID-19 sparked instructors' interest in ungrading as way to express their concern for students and show humanity amid the ongoing catastrophe, I then turn to analyzing the relationship between, on the one hand, key pedagogical values teachers hold and the affordances or possibilities they see in ungrading as a form of writing assessment. I illustrate this relationship through three examples — Carter, Sterling, and Taylor — showing how in each case, teachers' values predict or guide the affordances ungrading has for them. A teacher who prioritizes empowering students, for instance, is likely to pick up on ungrading as an antiracist practice, while a teacher who values their agency as a professional may perceive ungrading as validating their own labor. I then shift to the second half of this chapter, using data from the 2021-2022 set of interviews to construct a longitudinal study. As Sterling's and Taylor's experience shows, the longer instructors work with ungrading, the more impact it has, altering or deepening their values and commitments and (re)shaping the course in important new ways. Ultimately, the values that instructors bring to ungrading and the affordances they encounter within the institutional environment they inhabit work together, cultivating an assessment ecology that sustains (or not) instructor interest in alternative assessments and (re)shaping the writing classroom to make more educationally-fruitful, liberatory ways of being for both teachers and students.

Methods

I adopted what I would refer to as an intuitive process of coding for initial review of the data (2020-2021). I started by generating descriptive codes. I then drew on work such as *Coding Streams of Language* and Saldana's reference text (2016) to suggest further steps for (re)combination an analysis. In analyzing the data, I followed a cyclical process of performing initial, descriptive coding, followed by combining and synthesizing the codes into different dimensions that captured and aspect of participants' relationship with grades and/or with ungrading. I followed this cycle at least three times.

During the first round of coding, in October 2021, I created codes or dimensions that captured the advantages they perceive in ungrading ("Attends for Social Inequities — COVID"), then organized them under a larger, umbrella dimension termed "Affordances" — a term I chose to capture the opportunities that ungrading opens in the classroom or the conditions it facilitates, such as greater empowerment for students or a sense that their own labor as educators was more fully rewarded. Throughout, coding began with small, highly descriptive codes that drew on participants' language choices as much as possible (in vivo coding), to center their perspectives and experiences. I drew on observations and note-taking to combine the codes into larger themes. This round of coding ended with a list of affordances or outcomes that instructors at UTK perceived in the practice of ungrading.

Two subsequent rounds of coding were performed in February 2022, to draw out observed patterns in the data. First, I created a new dimension to capture the affordances of grading, to better understand how participants' impressions of ungrading may differ from prior experiences. At this time, I also separated the "affordances" dimension into "affordances" and "anti-affordances," later called "institutional affordances," to distinguish the advantages that participants see as inherent to ungrading and the institutional factors or conditions, which make it possible for them to explore it further.

At this time, I also noticed that participants' experiences with ungrading were often shaped by prior teaching-related values, or the conditions and moves that teachers prioritized and thought important in the classroom. I used values coding (Saldana 2009) to capture the beliefs, values, and goals that educators brought to their experience with ungrading; this dimension allowed me to recognize more directly that assessment is an inherently value-driven project (Huot 2002, Schwarz 2020). In lieu of conventional, one-paragraph memos (Saldana 2009), I preferred creating extensive, multi-page integrated memos that connected the codes with each other and relied on individual, constitutive quotations to develop and clarify the definition of each code. Finally, I compared the "values" code with the "affordances" code, at both the whole dataset and individual participant level, to explore how particular beliefs and goals predicted what participants found attractive or worthwhile in ungrading. The results of these comparisons appear in charts below. The comparison allowed me to treat values and affordances as "independent/dependent [variables]" (Deterding and Waters, 2021, p. 720) — in other words, to (re)consider the connections or influence, whether uni- or bi-directional, between these two codes. Still, this analysis would benefit from further study, especially attention to interrater reliability and member checking, as neither practice was possible for me at the time.

I listed the most relevant codes in Table 5.1 and Table 5.2. In Table 5.1, I list codes that capture the professional/pedagogical values instructors bring to ungrading; these values preexist instructors' use of ungrading and shape their experience(s) with it. I chose to describe values as verbs to capture teachers' vision for how they want their teaching practice to operate, or what they want their teaching to achieve.

In Table 5.2, I document the affordances, or opportunities, that ungrading as an assessment practice allows instructor. These too are phrased as verbs, to capture the active, potentially transformative role of ungrading in the classroom. The two columns, and the ordering of bullet points within the columns, is not intended to depict any particular relationship; I discuss the relationships among these codes and values more fully below. A full codebook is available in Appendix H.

I used these codes, and their interactions with each other, to develop the arguments in the first half of this chapter. For UTK writing instructors, the outcomes of ungrading were fluid, as the practice is (re)shaped by their own commitments. Some instructors stressed its value for one affordance; others emphasized a different affordance, depending on the variable commitments or values they brought to the work of teaching. Additionally, as I show in the second half of this chapter, the longer UTK educators used ungrading, the more the values they brought to the work developed and expanded, so that sustained use of alternative assessment practice made possible a more equitable, theoretically robust assessment approach.

Covid and ungrading

However, before I turn to the intersection between the affordances ungrading offered to instructors and their values, I want to discuss one significant feature of instructors' experience, especially during the first year (2020-2021) of this study: the role of COVID-19 in driving instructor uptake of ungrading. As I explained earlier, the COVID-19 pandemic, which shuttered universities during its first months and coincided with the resurgance of Black Lives Matter, at sharpened the need for educational and assessment practices that are equitable, accommodating, and flexible enough to support students through widespread social disruption. The early months of the pandemic — and instructors' goals for their teaching during them — played a key role in driving the growth of ungrading at UTK. Discussing this role is important to shedding light on the varied outcomes of ungrading and its correspondence with values beliefs about teaching, as teachers' interest in ungrading as a possible response to COVID-19 previews the relationship between teachers' values and ungrading's outcomes I discuss later in this chapter.

At UTK, COVID as a factor in instructors' decision to ungrade starts strong, then decreases over time, showing that as social exigencies changed, other motivations for ungrading rose to prominence. Table 5.3[83] documents the number of data segments that connect instructors' decision to ungrade to their desire to meet the challenges of the pandemic, and/or the increased social inequity produced by the pandemic. By sorting the data segments by the term during

[83] If participants spoke in both a focus group and a follow-up interview, the follow-up interviews are omitted from Table 5.3. (Follow-up interviews are omitted from analysis in the dissertation.)

Table 5.1 List of teachers' values and beliefs about writing/writing education, Fall 2020-Spring 2022

Code name	Definition
Prioritizes fairness	*Instructors value education* that avoids penalizing students and/or finds alternative ways to acknowledge student accomplishments outside of quality measurements. The teacher is no longer the judge or arbiter of quality. C.f. Elliot (2016).
Recognizes students' subjectivity	*Instructors value education that* acknowledges the place(s) students come from and their individual needs over objective, one-size-fits-all institutional expectations. Teachers have an obligation to fill the gap.
Comes alongside students as a mentor/guide, rather than a judge	*Instructors value education that* situates them in a non-hierarchical position relative to students, a mentor/coach or "guide on the side" versus an arbiter of grades. Instructors value being open and vulnerable with students.
Practices self-care	*Instructors value education that* acknowledges the stressors and mental health needs that students and teachers bring into the classroom with them and seeks to meet them if possible.
Empowers students	*Instructors value education that* ensures students have the knowledge and space/structure to make choices that let them grow as writers and reach their goals. Antiracism is coded here.
Centers rhetorically-informed writing education	*Instructors value education that* centers best practices and goals from rhetorical or writing studies scholarship, such as transfer, genre, reflection, and the distinction between higher and lower order concerns.
Practices sound pedagogy	*Instructors value education that* aligns with best practices in education. They value being effective teachers and receiving high-quality student work. This code encompasses teaching practices that "translate" across disciplines, e.g. it is not rhetorically specific.
Enjoys teacher agency	*Instructors value* the freedom to make decisions for themselves about how they do their work and/or how they spend their time. This extends to environmental factors (they value support from students or colleagues) and personal ones (they value feeling confident in their own decisions).

Table 5.2 List of teachers' perceptions of affordances of ungrading, Fall 2020-Spring 2022

Code name	Definition
Attends for social inequities outside of race	*Ungrading* acknowledges and/or removes social inequities from the assessment process, so that students are assessed only on relevant work for the course. Items may be coded here even if they do not mention ungrading explicitly, so long as participants speak directly about the relationship between social inequities and grades/grading.
Attends for social inequities — COVID	*Ungrading* acknowledges and accommodates the challenges that emerged or were amplified b/c of Covid, and that students or teachers faced. Covid may be glossed as "2020," "these times," etc. Items may be coded here that do not mention ungrading, so long as participants explicitly recognize the Covid-related challenges students face.
Develops student-teacher relationship	*Ungrading* fosters a non-hierarchical relationship between students / teachers, minimizing teachers' work as arbiter in favor of work as coach / mentor / guide.
Surfaces antiracist teaching practices	*Ungrading* cultivates opportunities for students with diverse raciolinguistic backgrounds to succeed and thrive. Teachers have the opportunity to pursue antiracist strategies explicitly.
Interacts with pedagogy	*Ungrading* supports those ways of being in the classroom, both students' ways of being and teachers' ways of being, that facilitate learning and writerly growth, e.g. lessened anxiety, risk-taking, a focus on transfer, etc. Items may be coded here if instructors discuss how ungrading constrains or limits student learning.
Values labor (instructor labor)	*Ungrading* makes it possible for teachers to spend their time and energy how they want, and/or to develop into the kind of teacher they want to be. Both quantitative (the amount of time) and qualitative (what goals or tasks the time is spent on). Discussion of how ungrading allows teachers to give feedback instead of grade/justify a grade is coded here.
Values labor (student labor)	*Ungrading* makes the course easier to pass, often via assessment design that removes barriers and allows student work to pay off in the form of success. Items may be coded here if instructors fear that ungrading undermined success.

Table 5.3 Instructor mentions of Covid as motivating their adoption of ungrading

Semester/Year Taught	Mentions of Covid
Taught Summer 2020 (2 participants)	14 (Average of 7 references / participant)
Taught Fall 2020 (2 participants)	7 (Average 3.5 references / participant)
Taught Spring 2021 (3 participants)	10 (Average 3.3 references / participant)

which instructors used ungrading (Summer 2020, Fall 2020, or Spring 2021), Table 5.3 indicates a declining interest in ungrading over the course of the study period.

To be clear, Table 5.3 documents only instructors' *explicit* discussion of COVID as a motivating factor for *taking up* ungrading, omitting attention to COVID as affecting instructors' experience of ungrading (for instance, teachers wondering whether COVID made ungrading more confusing for students). It also omits attention to broader discussions of social challenges students face, such as limited socioeconomic means or college preparation, which, as participants point out, are deeply intertwined with the pandemic. Thus, Table 5.3 likely underreports the number of times instructors consider or account for the pandemic in their ungrading practice. Still, conversations with participants in Summer 2020 are *dominated* by Covid. Kerry reports that "in general the world is so fucked up and unfair right now" that in the wake of the first COVID wave, the desire to reduce unfairness left little room for more targeted concerns about equity or pedagogy. Clearly, instructors come to ungrading with a deep desire to make things right, to the degree possible that summer, for their students and themselves.

Yet the choice to use ungrading, and perception of it as meeting the needs of the moment, hinges on other interrelated values and assumptions about what makes for good learning. Kerry, one of the teachers *most* concerned about COVID, says of their experience ungrading during the 2020-2021 academic year:

> [Using ungrading] just feels really good when — I mean, it probably feels good in normal times too — but it felt really good during COVID when people were like Oh my God, I have to miss a thing! Or I can't come to a conference with you, or whatever, and it's just like, Oh, really man? It's just not a big deal, I can make it up for you, I can work with you. If you want to meet with me next week or resubmit a homework assignment that you missed, I'm happy to take it. But also like? you took your mom to the hospital to get a COVID test, I'm not gonna be all up on your case about it.

It's clear that Kerry primarily sees ungrading as affording students the opportunity to care for themselves and family members, and teachers the opportunity to "feel really good" about their work. That Kerry values this flexibility as key to their teacherly identity, however, hinges on assumed values — the importance of prioritizing students as humans with complex lives and of mutual care. As significant a role as COVID plays in motivating instructors' uptake of ungrading, even this motivation is linked to and shaped by teacherly values that preexist and structure educators' use of ungrading. The interrelationship between teacher assumptions about foreshadows the point I make at greater length in the next section, that affordances shift and change depending on the values teachers bring to the work. Here, I simply want to observe that this interrelationship reflects the ecological nature of ungrading. In a landscape constricted by the pandemic, new ways of caring for each other — including new ways of assessing — emerged, shaped by more enduring commitments. As COVID increasingly took a back seat, its prominence as a key motivating factor fell away; and indeed, by the time I conducted the second round of interviews (2021-2022), I found that my interview question about COVID was often the first mention of the disease in our entire conversation. COVID receded (for better or worse) from the educational landscape, and other values emerged, to meet new exigence and teacherly

motivations. In the next section, then, I turn to exploring how UTK instructors' values shaped their interest in ungrading and their perception of its value, for reasons that extend beyond its potential to meet COVID exigences. I also discuss how their use of ungrading, in turn, worked backwards to (re)shape the values instructors bring to the classroom.

Correspondence between ungrading and teacher values

As I described earlier, my central claim in the chapter is this: that for UTK instructors, the affordances they discover in ungrading correspond with their preexisting values and at times (re)shape those values. Inoue asserts that as a key part in the assessment ecology, alternative practices such as labor-based grading flow outwards and reshape the classroom landscape in ways that blunt the impact of discrimination. I argue that the direction can work the other way as well, with educators' commitments flowing backwards and reshaping the use of alternative assessment practices in ways that change the affordances those practices offer.

Table 5.4 documents the correspondence between the values instructors bring to the classroom (the Y axis) and the affordances they recognize in ungrading (the X axis)[84]. On the left hand side of Table 5.4 (the Y axis) I list instructor values, along with the total number (frequencies) of data segments in the initial focus groups or interviews coded at that value. (This number is more than the sum of the numbers in the associated row, because some data segments coded at that value do not mention a particular affordance or advantage of ungrading; for instance, not all of the 66 segments coded at *fairness* are also coded at any particular affordance; some segments are not coded for an affordance at all.) Along the top of Table 5.4, or the X axis, I list the major affordances of ungrading, meaning the advantages or outcomes UTK instructors perceive in their practice. These affordances are primarily framed as positive by instructors. As I discuss in the pages below for instance, educators speak positively of ungrading's ability to work towards antiracist ends (Sterling) or validate instructor labor (Taylor).[85] Cells that represent a convergence between a value and an affordance are marked with a double star (**) to draw attention to the way that instructors' values shape what they perceive as the chief affordances of ungrading. The coding schema which grounds Table 5.4 and includes a complete definition of each code and examples may be found in Appendix H.[86] [87]

[84] Both Broad (2003), Schwarz (2020), and to a lesser degree Huot (2002), discuss the importance of value in assessment, the way value drives assessment. I thought of this argument in constructing my coding around value.
[85] Not *all* affordances are positive, however, as for instance instructors also speak of ungrading (re)shaping their pedagogy in ways that deny them previously available teaching strategies, such as the ability to incentivize certain behaviors (Jamie). Further exploration of this distinction, however, is beyond the scope of this chapter, which primarily seeks to document the relationship between instructor values and the affordances noticed in ungrading.
[86] *Pedagogy* in this dataset is both an affordance and a value. As an affordance, along the *X* axis, *pedagogy* references the ways ungrading supports a wide array of educational practices that contribute to learning and growth; as a value, along the Y axis, *pedagogy* references UTK instructors' belief in the importance of best practices in their work and commitment to being effective educators. There is naturally overlap between *pedagogy* the value and *pedagogy* the affordance, since valuing education means UTK instructors are likely to appreciate ungrading's impact on their teaching practices. These are not identical sets, however, and the overlap between pedagogy the affordance and other values points towards a complex understanding of good writing, as I describe below.
[87] *Pedagogy* the affordance also includes "develops the student-teacher relationship." for the sake of space. (Of all segments coded as the value "mentoring students," 14 are coded under "develops student-teacher relationship.)

Table 5.4 No. of data segments coded at particular values versus at ungrading affordances

	Affordances						
Values		Soc. Inequity	(Covid)	Antiracism	Pedagogy*	Instructor labor	Student labor
	Fairness (66)	11	4	4	6	2	20**
	Subjectivity (28)	2	16**	1	4	1	0
	Mentor (23)	0	0	0	15**	0	0
	Self-care (15)	0	2	0	4	9**	0
	Empower (33)	0	1	7	17**	0	1
	Rhet. Eff. Education (14)	0	1	1	9**	2	0
	Pedagogy (28)	0	0	1	11**	9	0
	Teacher Agency (53)	0	1	1	1	23**	0

What Table 5.4 demonstrates, first, is the correspondence between particular values and related affordances, suggesting that UTK instructors' deep-seated ideas about how (writing) education should work and who they want to be in the classroom directs their attention to certain affordances in ungrading. Caring about their own agency in writing education, for instance, leads instructors to appreciate the way ungrading values instructors' labor in the classroom.

Their concerns about recognizing the subjectivity of students correlates with their attention to ungrading's responsiveness to the increased inequity created by Covid. This correspondence does not mean that instructors' values and the affordances of ungrading are identical; that ungrading allows for greater responsiveness to the challenges of Covid is not simply another way to say that the instructors using ungrading care about the subjectivity of individual students, or teachers. Rather, the correspondence calls attention to the way that ungrading makes it possible for teachers to enact the values and ways of being they bring to the classroom but have been unable to realize until this point.

So when instructors speak of how ungrading meets the exigencies of Covid, that line of argument assumes both the value of recognizing the inherently varied experiences that shape students' performance in the writing classroom and also the difficulty of acknowledging that diversity within a conventional system. These values are perhaps more fully articulated through instructors' experiences with ungrading, brought into light and strengthened; but they are not identical with the outcomes people recognize in ungrading. Put another way, instructors recognize the affordances they do in ungrading because it reshapes the landscape of their classes in ways that make possible desired ways of being and teaching. Kerry, for example, speaks of seeing "students go through some crazy stuff this [term], like working full time, caring for family members with Covid," then suggests, "one thing I took away was just like Oh! helping people navigate life on top of school makes a better class environment, not a less rigorous or less serious about learning class environment." Coded both as an affordance of ungrading (accounting for Covid) and as a value, of prioritizing students' subjectivity, this passage speaks to Kerry's appreciation for the way ungrading made it possible to meet pandemic exigencies and assumes the importance of putting students' individual, fluid needs above more static, institutional expectations such as rigor. This commitment to students is at once a given for Kerry, preceding their sense of the affordances of ungrading in relation to Covid, and also articulated through those affordances, something that Kerry "takes away" from their experiences observing and teaching students during a Covid-impacted semester.

Second, Table 5.4 shows that multiple values coalesce around pedagogy as an affordance, or ungrading's impact on teachers' pedagogical decisions. This convergence suggests that UTK instructors have multiple, complex, and overlapping values that they bring to the work of teaching, and that undergird their appreciation for ungrading as a pedagogical choice. Ungrading, as I show elsewhere, is often received by UTK instructors as a solution *specifically* to the problem of grading, a time-consuming, arbitrary (River) endeavor that demands instructors justify grades to students rather than help them learn (Jamie). But ungrading is not *only* a grading solution. Rather, ungrading for UTK instructors during AY 2020-2021 functions as a catalyst, helping to bring into being the particular *kind* of educational space desired — one in which they can mentor students, support student empowerment, and effectively teach rhetorical practices. A

non-hierarchical, student-centered, and rhetorically-informed approach to education becomes (more) possible within the landscape of an ungraded classroom, than a conventionally graded one. This is not to say that ungrading always has a positive impact; the term for this affordance – – ungrading's interaction with pedagogy — gestures towards the way that a landscape where standard measurements no longer exist may constrain workable education as well as allowing it to flourish. The point more precisely is that ungrading interacts with teachers' values in ways that (re)shape the landscapes or space of the classroom, altering relationships to make possible a pedagogy "with the grain" of teachers' values.

Illustrating the convergence between values and ungrading affordances — three examples

In the next section, I turn from exploring the convergence between values/affordances overall to three specific narratives, illustrating how individual teachers find that ungrading's outcomes correspond with their own values or goals in taking the practice up. Because the teacherly narratives I select emphasize different outcomes, or affordances, of ungrading, this section also demonstrates the fluidity of ungrading as a practice, that it achieves different things depending on who is using it and for what purpose. I focus on three specific instructors: 1) Carter, whose interest in fairness overlaps with ungrading's potential to recognize student labor; 2) Sterling, whose commitment to empowering students overlaps with ungrading's impact on pedagogy; and 3) Taylor, whose interest in professional agency corresponds with ungrading's potential to validate teacher labor. This variation — especially combined with the longitudinal data in the second half of this chapter, which return to Taylor and Sterling — allows the kind of comparison-and-contrast that makes for effective case study. In the space below, I include a brief description of the instructor, Table 5.5 documenting the intersections between their own values and the affordances they perceive in ungrading, and an analysis of the role their values played in (re)shaping the assessment landscape within their course.

My decision to foreground particular instructor stories, rather than proceeding affordance-by-affordance, is informed by my use of case study. Conclusions that are generalizable to the wider population, not only UTK but writing programs and universities at large, is not possible within case studies, which as my methods chapter discussed, works towards "concrete, context-dependent knowledge" (Flybjerg p. 224, 2006). This is not to say that my points here have no relevance for other institutions. Rather, by focusing on particular instructors' encounters with ungrading, I am able to shed light more broadly on the turn towards alternative methods of assessment at UTK in the early 2020s — and then, through cautious connection-making from example to example, develop sufficiently thick description to make that turn relevant to other writing programs and universities.

Carter

Carter's example illustrates the close link between a teacher's values related to teaching writing and their perception of ungrading's affordances, given that Carter — concerned about unfairness in grades/grading, especially during COVID — finds alternative forms of assessment to be fairer and therefore stronger pedagogy. At the time of the study, Carter, a GTA, had fewer than three years' experience teaching first-year writing. Reflecting on their experience teaching during the

Table 5.5 Values, listed by prevalence — Carter, 2020-2021

Values, listed by prevalence (Carter)	No. References
Fairness	20
Teacher agency	6
Recognizes students' humanity	4
Empowering students	2
Comes alongside students	1
Rhetorical effectiveness	1
Strong pedagogy	1
Self-care	0

pandemic, Carter notes the challenge of fairness not only during the immediate upheaval of AY 2020-2021 but the variable impacts on students across socioeconomic class and racial groups, observing (for instance) that students raised in wealthier homes were able to continue their studies online and enter college that fall much more prepared for their coursework than classmates with fewer resources. The importance Carter places on fairness appears clearly reflected in Table 5.5, in which twenty of the passages coded for values reflect Carter's emphasis on fairness. Carter, in reflecting on the affordances or outcomes of ungrading in their classroom, prioritizes the practice's ability to improve their pedagogy and correct for the social inequities of COVID (see Table 5.6). These affordances do not, of course, correspond perfectly with Carter's commitment to fairness: "interacts with pedagogy" in particular is an expansive code, gesturing towards a range of improvements or changes that ungrading makes to teaching practices[88]. I focus on Carter, nevertheless, because their belief that conventional grading is unfair, particularly during the pandemic, precedes their perception of ungrading as a fairer, more equitable approach, particularly during COVID; this in turn primes Carter to recognize *other* wasy that ungrading may allow for greater fairness in their own pedagogy, reconstructing the classroom to treat all students equally. Put more simply, Carter's experience with ungrading aptly illustrates that interest in the practice is driven by preexisting values and that experience with ungrading then (re)shapes and expands those values in productive ways for instructors' pedagogy or approach to their work[89]. See Table 5.6.

Key to Carter's experience is their increasingly strong belief that conventional ungrading is unfair in a way that undermines or weakens effective teaching. For Carter, grading required accurately measuring the quality of student writing, then punishing students who fell short. Once COVID hit, this task became impossible for Carter, both logistically — how does one calculate grades amid a global pandemic — but also emotionally and professionally. Grading was not *part* of Carter's work as an instructor; rather, it *interfered* with their work. Carter relates:

> I was very stressed using the traditional grading scale because of the way it has the numbers, like a lot of them have okay, you get an A if you have — if you do 14 out of 17 points in this category, you get an A, or whatever; and I always had a really hard time quantifying what number I would give someone, because usually it's a range, you know, I could never figure out how to pair their fluency or their sentence structure with a number, like that didn't really make sense to me, because that seems just really subjective. And so a lot of times I would feel bad after grading, because I would be like I don't know if I was completely fair, I don't know if I was entirely objective with my grading, and sometimes that's why I would add a few grace points here and there because

[88] In Table 5.6, "interacts with pedagogy" includes 1 mention of developing the student-teacher relationship

[89] At a global or aggregate level, of course, values around fairness correspond most closely with ungrading's ability to validate student labor. Yet of course, aggregated data masks nuances and variances in individual cases, such as here; and in fact, Carter's case illustrates another feature of the overall table, that fairness corresponds *secondarily* with ungrading's ability to account for social inequities, both generally and more particularly during Covid. This correspondence reinforces that the relationship between values and particular affordances is not one with statistical significance, only a locally-notable pattern that underscores the importance of teacherly values in shaping our ungrading practice(s).

Table 5.6 Perceived affordances of ungrading, listed by prevalence — Carter, 2020-2021

Affordances, listed by prevalence (Carter)	No. References
Interacts with pedagogy	9
Attends for social inequities — COVID	7
Antiracism	6
Instructor labor	4
Student labor	2
Attends for social inequities outside of race	1

> I never knew if I was being genuinely fair to the students, and that caused me a lot of stress and a lot of thinking about it after grading. It was just really hard to try to be a coach for [students] almost, like a writing coach, but also a punitive person.

Carter's visceral reaction to the unfairness of grades and grading is crucial. They describe grading as inherently paradoxical, on the one hand meant to be objective, yet on the other, inevitably *subjective*, as instructors attempt to pin a number on students' "fluency or sentence structure." The only way out, is to subvert the grading process entirely, "add[ing] a few grace points here and there" to make up for inevitable unfairness, particularly during COVID. Indeed, much of Carter's frustration with grading is tied to the pandemic. Carter explains: "I feel so bad, negatively evaluating them right now [during Covid] and making their life so much worse, and so I ended up adding five points, like five grace points back, after I graded based on the traditional rubric, and then I added five points, so if they got an 83, they would end up with an 88 or something." Acknowledging the impossibility of being completely fair with students, particularly during a global crisis, Carter adds in grace points for students, effectively undoing the claim grading makes to fairness via an accurate, objective read of students' work.

The inherent unfairness of grading, in fact, interferes with Carter's vision for what it means to be a good teacher, effectively supporting and coaching students through the writing process. In my archival chapter, I argued that historically, grading functions as a mechanism to hold students accountable to a Platonic ideal of good writing, regardless of their own background. The reaction Carter describes to grading suggests instructors experience grading as having that same function, to reinforce the Platonic ideal regardless of social or personal contexts. Carter, already concerned that grading demands objectivity for inescapably subjective judgments, hates imposing those judgments on students amid a global catastrophe that deepens social inequity. That Carter does not want to "make students' lives worse" acknowledges the highly contextual, personal situations that students face and implies that what makes grading bad is its refusal to make that acknowledgement, resulting in unfairness. The only solution is to go outside the system altogether, ignoring how students measure up to the Platonic standards and instead adding "grace points" that have nothing whatsoever to do with students' perceived abilities. In the end, grading makes it impossible for Carter to do their job. As Carter explains, it's not possible "to be a coach for students" while also being a "punitive person," always holding students to a Platonic ideal and punishing deviation by subtracting points. Grading and teaching conflict with each other, causing frustration for teachers who value their ability to be fair, supportive educators in the classroom.

Carter finds ungrading, in contrast, fairer, since teachers are no longer obligated to measure the (perceived) quality of student work against a singular standard. This in turn makes it easier for Carter to accommodate students' needs during the pandemic — and, ultimately, to make their teaching practice fairer *outside* of pandemic exigencies as well. Comparing their ungraded semester during 2020-2021 with previous, conventionally graded semesters of teaching, Carter says: "this year was so much better because I knew that the model I had was fair, and I didn't have to think about it later and dwell on it, because I was like, okay, I gave them a grade, this was a grade they deserved because of their effort in the class and now I'm going to move on from that and not think about it anymore." As Carter describes above, conventional grading was

difficult, particularly during COVID since it required teachers to quantify exactly how good student writing. Ungrading, however, emphasizing effort instead of (perceived) quality or alignment with a Platonic ideal, making the process of arriving at a grade clearer. The tension between the necessity that grades be objective and the inevitable subjectivity of taste, in assessing the quality of student work, goes away, enabling Carter to carry out their work productively. This in turn makes it easier for Carter to assess students' grade fairly, cushioning the otherwise substantial impact of social inequity on grades during the pandemic. As Carter explains, even though wealthier students had the technological resources to focus on learning amid the pandemic, ungrading ensured that students "weren't going to be disadvantaged for [something] they couldn't control;" it "levelled the playing field" and "made [the] class a lot more equitable and better overall for students." Carter's rubrics for each paper list efforts, or labor, such as participating in peer review or including the correct number of sources, tasks which students who come from a poorer background may complete regardless of how much prior training and education in writing they have. Because students' labor is less closely tied to their socioeconomic status, for Carter, the grade depends less on students' background or other circumstances out of students' hands and in turn provides a fairer, more equitable way to assess work, even as the pandemic brings social inequity into sharper relief.

Though Carter's original concerns about fairness are rooted in the experience of trying to assess work amid a pandemic, Carter also comes to value other ways ungrading ensures fairness, among them students' political stances and raciolinguistic backgrounds. This in turn improves Carter's pedagogy, ensuring they feel like a more successful teacher in the writing classroom. Carter notes: "There were a couple of topics that I wouldn't have agreed with the student about, but because of contract grading, I didn't have that ability to be subjective, whereas I feel like with a traditional grading rubric, there's some room for subjectivity, cause I could have made their score 12 out of 17 for content or something, if I didn't like their paper." Acknowledging the tension between (again) the need to grade objectively[90] and the difficulty of being objective about some subjects in particular, Carter puts a finger on an advantage of ungrading, that in shifting assessment to effort or labor instead of a subjective assessment of quality, it allows teachers to respond more fairly to student work. For Carter, there is no need to worry about potentially penalizing student work because they "didn't like their paper" topic. Ungrading makes it possible for Carter to concentrate on whether students "complete[d] the assignment," resulting in a grade that is more transparently fair. Carter also points to students' raciolinguistic background as an area where the greater fairness of ungrading is valuable. Students with "English speaking parents, [who are] wealthy [or who] come from a private school [which] could afford to teach them standard American English" are no longer automatically privileged within an ungraded system. This in turn leaves room for multilingual students or students from socioeconomically disadvantaged areas to succeed as well. Ultimately, by shifting to a grading practice that prioritized students' effort or labor in the course over the perceived quality of their

[90] I do want to be clear that in some cases, accounting for the positions that students take up in their papers is in fact the more ethical, equitable response (if not precisely what we mean when we think of *fairness* or objectivity). Students who write discriminatory work, for instance, may do harm to their classmates and educators, and the impact of their writing should not be ignored (Patterson, 2015). Often, however, students' stances are distasteful without being discriminatory, and teachers feel obligated to do right by students despite disagreement with their work.

work, Carter feels as though they can more more effectively reach their goals as a teacher: correct for the unfairnesses of the pandemic and treat students more evenly in their classroom, despite variance caused by ongoing social crises, political difference, or linguistic backgrounds.

Nevertheless, Carter remains sensitive to the way that unfairness persists even within an ungraded system. While shifting the focus from *quality* to *labor* reduces the gap, it does not close altogether. As Carter explains, "some [students] tried a lot harder than others" but ended up with the same or even a lower score, likely (Carter hypothesizes) based on the way that their own ungrading approach interacted with Canvas, the learning management platform that UTK instructors used to track and record grades. As Canvas is the primary way that instructors organize assessment records and communicate assessment to students, the interaction between Canvas and a particular assessment approach can affect and even undermine the desired outcomes of that assessment approach, as Carter describes here[91]. This acknowledgement is key, since it underscores (again) that ungrading is not a magic pill. Most instructors, including Carter, find that ungrading addresses questions of fairness in the ways that counts; as Carter explains, it is much easier for *them* as a teacher to do right by their students, sidelining questions such as stylistic tastes or even political beliefs, so that the writing construct being assessed remains as clearly defined and uncontaminated as possible. Yet as an ecology, the assessment process includes but is not limited to the teacher; it also interacts with and is (re)shaped by the LMS, historical or conventional practices, and students' own experiences. As Carter points out, reflecting on their ungrading design, especially as that design was constructed within the LMS (Canvas), ungrading may *become* unfair at other points or areas within the larger ecology, even if the instructor is better and more fairly able to respond to student work.[92]

Carter's example is significant for two reasons. First, crucially, Carter is not alone in valuing fairness. Among the values that UTK instructors bring to their work teaching first-year writing, fairness is by far the most common (66 instances). Fairness is the reason instructors appreciate how ungrading interacts with and rewards student labor, though also for instructors' appreciation of ungrading's ability to soften social inequities. Carter's experience of ungrading is not perfectly aligned with other instructors' experiences, in that for Carter, the importance they place on fairness drives their attention to ungrading's ability to alleviate pandemic inequities and improve their own teaching practices, versus (as is more common for instructors broadly) adequately reward students' labor. Yet Carter's story captures a common experience among

[91] While instructors' experiences with Canvas are outside the scope of this project, it's worth noting that most folks expressed frustration and difficulty getting ungrading to work within Canvas. Having used Canvas myself, I suspect that the platform is oriented around points-and-percentage based forms of conventional grading as the default and as a consequence is resistant to other, more liberatory approaches.

[92] For example, scholars such as Kryger and Zimmerman (2020) and Carillo (2021) have expressed concern that some versions of labor-based grading, particularly those which rely on labor logs, may disadvantage neurodivergent students, who may need more time to complete the same tasks or work in a way that is not easily trackable. Though Inoue nods to the complexities and variation within human experience, his model, as Carillo and others point out, assumes that, in his words, "everybody has the same 24 hours in a day" (p. 84, 2017). What Carter's experience gestures towards is the need for fuller and perhaps more nuanced understanding of fairness, one that not only works after-the-fact to work with individual students but starts with the assumption that students' individual needs need accommodation and flexibility. See also Elliot's description of fairness (2016), along with Carillo's (2021) analysis of the need for systems that to be accessible from the start.

UTK instructors, the frustration with the requirement grading imposes on them, that they hold students to the old Platonic ideal of writing and punish those students who fall short. However dominant that ideal in the early years of the program, UTK instructors no longer want to be the "final arbiter" or authority about what counts as good writing, as Jamie puts it. This is especially true during COVID, as instructors, like Carter, are acutely aware of how the ongoing trauma of the pandemic distorts students' abilities and make arbitration about writing quality invalid. By replacing these subjective questions with (more) objective ones about students' effort and their completion of tasks, ungrading gives UTK instructors relief from the pressures and perceived *un*fairnesses of conventional grading and lets them proceed in a way more aligned with their values. Instructors are able to give over punishing students and instead celebrate and support their growth as writers.

The importance of fairness is further illustrative here, since my impression is that Carter does not take up ungrading and find it to be fair but rather values fairness and takes up ungrading, hoping it will meet that value.[93] Among UTK instructors, consistently, the affordances of ungrading are perceived as such not because of any particular virtue in ungrading but because of the values that instructors bring to the work. This is not to say that ungrading is altogether morally neutral. At the very least, it offers an alternative to grading students against a (white supremacist) Platonic ideal. Ungrading is, simply, dynamic and personal, its affordances emerging from the particular classroom ecologies that instructors create with their students — in Carter's case, the exigencies of responding to COVID and the need to tread lightly around political topics. The value that Carter places on fairness throws into relief ungrading's ability to meet these needs, while other UTK instructors value other things and so perceive different affordances as most key. Ungrading, however, does not altogether satisfy Carter's desire for fairness, as Carter finds its treatment of student effort uneven. Ungrading's failure here is key, as it sets up a contrast with other UTK instructors who not only find that ungrading fully answers their deep pedagogical values but transforms them. I discuss one such case below: Sterling, who finds their use of ungrading to afford students more agency challenges them to reflect and deepen their commitment to empowerment and antiracism in the classroom.

Sterling

A GTA relatively new to teaching, Sterling values pedagogical practices that empower students – – in other words, that affirm students' agency and autonomy as writers. Accordingly, their experience with ungrading emphasizes its affordance to hone or sharpen pedagogical practices, as Sterling is, via ungrading, able to reach prior goals of student empowerment and agency. I selected Sterling as an example because these differences — especially when set against the similarities between Sterling's and Carter's teaching background and status as GTAs — make it

[93] To be fair, when discussing teachers' values and the (perceived) affordances of ungrading, there is a bit of a chicken-and-the-egg question. Perhaps Carter *does* come to value fairness through their experience with ungrading! Moreover, as I discuss in the second half of this chapter, instructors' experiences using ungrading for multiple semesters seem to suggest that teacherly values about what is important in writing, and in writing education, can be rewritten, deepened, and complicated. It is no good pretending that values are an entirely independent, static feature of the assessment landscape. My point here, however, is to suggest that instructors find ungrading attractive thanks to commitments that are not inherently connected to ungrading; Carter can value fairness without ever having tried alternative forms of assessment.

possible for me to illustrate how ungrading answers and even expands teachers' goals within the classroom. As my methods chapter discusses, exploring differences among the attributes that characterize individual experiences within a given case is key in building up a complex, nuanced picture of the whole.

Table 5.7 demonstrates Sterling brings to the work of ungrading a keen interest in empowering students and expressing their own agency and authority in the classroom, by making decisions they believe are pedagogically sound:

Correspondingly, Sterling finds ungrading most helpful in rearranging the classroom to enable more effective teaching and validate their own labor and identity as a teacher, as Table 5.8 indicates:

Where Carter's experience of ungrading was colored by COVID, Sterling perceives ungrading as meeting exigencies outside the immediate ones of the pandemic, as Table 5.8 illustrates. COVID plays a small role in Sterling's use of ungrading, yet their primary interest is ungrading's capacity to support student agency as writers. Sterling is driven by a deep commitment to workable, student-centered pedagogy and takes up and maintains their use of ungrading because they see it as allowing them to meet those goals.

One of Sterling's most important goals as a writing teacher is empowering students. Sterling is clear that what is "really important to" them as a teacher "is helping students identify as writers" Sterling explains:

> When I first learned about ungrading and what ungrading practices looked like, it became for me the way to achieve what I wanted for my students, in terms of developing as writers, see[ing] themselves as writers, also becom[ing] more critical and more aware in their reading and writing as they're moving on to go through the rest of their trajectory at the university.

Key here is Sterling's original goal in teaching, that students "see themselves as writers and think "critically" about how their reading and writing practices may be judged across their university education. Elsewhere in our conversation, Sterling contrasts this development with students "doing what they have to get the A" or "writ[ing] what the teacher wants them to write," performances which assume a singular standard of good writing, one by which teachers can accurately judge student work (Inoue). In contrast, Sterling values writing education which unsettles notions of an easily-judged standard and equips students with the metacognitive resources to perceive how writing is judged and make decisions for themselves about their writing practices.

This value precedes and shapes Sterling's use of ungrading, importantly. Ungrading does not introduce student empowerment as a new thing but in fact (re)arranges the classroom space in ways that make student empowerment feel to Sterling like an achieveable part of their pedagogy:

Table 5.7 Values, listed by prevalence — Sterling, 2020-2021

Values, listed by prevalence	No. references
Empowering students	21
Teacher agency	15
Rhetorical effectiveness	9
Strong pedagogy	9
Recognizes students' humanity	6
Fairness	3
Self-care	3
Comes alongside students	1

Table 5.8 Affordances, listed by prevalence — Sterling, 2020-2021

Affordances, listed by prevalence	No. references
Interacts with pedagogy*	26
Instructor labor	20
Attends for social inequities — COVID	5
Antiracism	5
Student labor	3
Attends for social inequities outside of race	0
Develops student-teacher relationship	0

> The things that I was doing for ungrading like using a grading matrix, doing reflective self-assessment, collaborative rubric development, all of those things I think helped students begin to see themselves as writers more and see their agency as writers and so the ungrading practices like became the means to an end to help achieve some of those things that I wanted students to leave my class with, like with knowing or with having.

I want to make a fine distinction here. On the one hand, the "things [Sterling is] doing for ungrading" have tangible effects in the classroom, helping students "see their agency as writers" and build confidence in their skills. On the other, Sterling's framing, that ungrading is a "means to an end," stresses that one reason ungrading *has* these effects is not (only) because of anything inherent to ungrading as a practice but because Sterling comes into the work of ungrading with strong prior values, values that center students' agency, writerly identity, and empowerment as writers. That ungrading does not, in itself, have these values but rather surfaces the values Sterling brings to the work is apparent through contrast with Carter, who also required students to submit reflections with projects (a version of self-assessment) but does not speak at length about student empowerment. For Carter, focused on different, equally legitimate, concerns of fairness and equity, ungrading has other affordances. For Sterling, concerned about student agency, ungrading and its constituent parts plays a significant role in (re)shaping their teacherly identity and practice to more fully support students' empowerment as writers.

A core part of Sterling's identity as a teacher is engaging with antiracist ideas. As Table 5.8 indicates, antiracism is not among the chief affordances Sterling perceives in ungrading. In fact, however, Sterling's attention to empowering students is expansive enough to include antiracism, which for Sterling centers the ways that students from raciolinguistically diverse backgrounds may benefit from a clearer sense of how their writing is being judged within certain spaces and how to make choices in light of those judgments[94]. Sterling explains:

> When I think about the way that ungrading is supporting or enacting antiracism, [that] goes back to helping students see, not just in this class but in every class they take, that their writing is being judged through a certain set of conditions and giving them the tools to see what those conditions are and understanding the way that they're producing writing in relation to those constructs within larger systems.

Sterling's mental model of antiracist language education here draws from *Antiracist Writing Ecologies* (2015), which Sterling credits as their "theoretical inspiration." In the text, Inoue posits that student work is "being judged through" the lens of a "white racial habitus," or (put another way) the standardized American English established as the Platonic ideal of writing; antiracist education consists, in part, of making this judgment visible to students. Sterling's description of antiracism here echoes their broader comments about empowering students, or recognizing student agency, in the emphasis on enabling students to make informed choices

[94] I coded passages that spoke broadly of making systems and standards of judgment visibly broadly as interacting with pedagogy, taking care not to overcode Sterling's references to antiracism. Nevertheless, as passages like the one on this page ("What I want to do is — and, I, this kind of goes back") indicate, Sterling's attention to empowering students as a key part of their pedagogy is key to the antiracist elements or commitments in their pedagogy, and so, Table 5.9 likely understates the depth and breadth of Sterling's commitment to antiracism.

about their language use, depending on the contexts they inhabit; their attention to the "conditions" used to judge student work, and the "tools" needed to interpret those conditions, picks up on Inoue's argument that writing evaluation is typically made from the perspective of whiteness, and normalized or standardized white language practices. "Empowering students," then, is for Sterling a complex value and produces complex outcomes in their pedagogical choices, at once emerging from their commitment to developing student agency and calling them back towards direct, explicit conversations that challenge student assumptions about what counts as good, academic English.

Sterling, looking forward to future semesters using ungrading, explains the importance of antiracist action to their pedagogical goals and identity:

> What I want to do is — and I, this kind of goes back to what I was saying about linguistic diversity, but I think having more conversations with students about antiracism explicitly, about language ideology explicitly, about power structures in institutions explicitly, and the way that ungrading is kind of fitting into that is something that I didn't foreground as much as I want to moving forward … I think ungrading opens that door in a way that wouldn't be there in the traditional composition classroom and so I think if I'm going to do ungrading, and I am, then I'm also taking that responsibility I think to engage students more in these like more complicated reasons why I'm doing ungrading.

Again, Sterling's attention to the connections among language and (institutional) power ground their interest in ungrading — the purpose they use ungrading for, in the classroom — firmly in both student empowerment broadly and antiracism specifically. That Sterling looks forward to future years ungrading indicates an interest in pursuing not only the practice but the concept of antiracist education more fully, leaning into ungrading's potential to transform pedagogical values and commitments; Sterling's interest in hosting direct conversations about language and power may point towards a sharpening of the values that made ungrading attractive in the first place. The longitudinal section below shows that Sterling actually goes on to have the conversations described here, so that where Carter finds ungrading partially but does not fully meet their prior values, Sterling finds that ungrading deepens their values.

This last point is key. I fear that in framing teachers' values as preexisting and shaping their use of ungrading, I make the process I am trying to describe — the impact of teacherly values and commitments on ungrading's outcomes — seem too unidirectional. This is not the case. In fact, ungrading may also work backwards and impact educators' values. I explore this aspect of participants' experiences with ungrading in the section below, using longitudinal data from the 2021-2022 round of this study, but I want to note that Sterling, in describing their use of ungrading to work for student empowerment, acknowledges ungrading's effect on their teaching and teacherly identity. Sterling explains: "Teaching for me has been such a space of construction of identity, becoming who I am as a teacher and how intimately my assessment is tied to that is something I never could have anticipated." Sterling clearly brings a strong sense of identity and values into the classroom, but the classroom — and the assessment choices that shaped it — exert pressure on that identity, (re)shaping it in powerful ways. As Sterling explains, "being the best instructor that I [can] be" includes (among other things) "helping their development as

writers in the way that I want to help them." Thus, while ungrading makes it possible for Sterling to realize their identity and commitments, the fact that teaching, and ungrading, also constructs Sterling's identity complicates the process. Though perhaps especially pronounced for a newer teacher like Sterling, more experienced educators also find their identity and commitments challenged through the use of ungrading, as the longitudinal section below shows. That Sterling changes as an instructor in response to taking up ungrading is encouraging, suggesting that if teachers are willing to lean into ungrading, to follow where it leads in their classroom, they may find it sparks new ways to think not only about assessment but about teacherly commitments broadly.

While I explore ungrading as transformative below, I want to briefly note that Sterling's expanding interest in the antiracist potential sheds light on the dispute over whether ungrading is, in fact, more racially equitable than conventional grading (Carillo 2021). Sterling's example underscores that ungrading does not so much flip a switch in the classroom as invite educators into a cyclical, ongoing process of identifying and taking up justice-oriented obligations[95]. Though programs may encourage and facilitate the use of ungrading, requiring or imposing ungrading as an institutional solution for racial inequity is of limited value[96]. Ungrading is not a magic pill; its outcomes depend on teachers' prior values and commitments. For Sterling, ungrading has (possible) antiracist outcomes, because antiracism is something that Sterling cares about. Instructors who do not care about antiracist education may not experience ungrading as a more equitable practice. Ungrading may work together with instructors' (or perhaps program's) goals to construct a more equitable landscape, it may invite instructors (and perhaps programs) further into that work, but without buy-in from instructors, it cannot accomplish this work on its own.

Taylor

Crucially, both Carter and Sterling center values that are strongly pedagogical — achieving fairness in assessment, or cultivating an antiracist classroom space. For my final example, then, I turn to Taylor, an experienced educator and NTT lecturer. As Taylor's use of ungrading shows, its outcomes are equally influenced by values that center teachers, as well as students — in this case, values oriented around teachers' own professional identity and agency. This outcome is important because, in the context of this chapter's argument that teachers' values shape the outcomes of ungrading, it shows how ungrading may be used to benefit teachers, as much as students. This observation challenges the exclusivity of ungrading scholarship's focus on validating student labor, extending that to teacher labor as well. Taylor, an NTT lecturer, comes to ungrading with a keen appreciation not only for students' learning and well-being but also for

[95] See also Le Guin, *The Dispossessed*: "There was process: process was all. You could go in a promising direction or you could go wrong, but you did not set out with the expectation of ever stopping anywhere. All responsibilities, all commitments thus understood took on substance and duration."

[96] This is not to say that writing programs should *not* impose more equitable assessment practices. Within the classroom ecology, as Inoue points out (2015), approaches such as labor-based grading can blunt the impact of biased judgments on the part of teachers. Some ungrading is better than no ungrading. At the same time, the danger is that if programs see ungrading as a magic pill, they may neglect the ways that such a requirement can also *create* more inequity, especially among faculty and students from minoritized populations (Craig 2020). For more discussion of institutional responsibility in ungrading, especially antiracist ungrading, see the previous chapter.

their own agency as an instructor, and as I discuss below, finds that ungrading validates their labor, and makes their work more professionally rewarding.

Table 5.9 below indicates that Taylor cares deeply about their agency as a professional and about fairness or equity in teaching:

Unsurprisingly, given Taylor's values, they prize ungrading for its ability to validate their own labor in the classroom and to attend to social inequities outside of race, especially those sharpened or brought into relief during COVID — outcomes which strongly correspond with a concern for fairness. Table 5.10 indicates Taylor's interests in ungrading:

Table 5.9 and Table 5.10 suggest a relationship between the value Taylor places on agency and autonomy as an educator, and on the right-hand side, Taylor's perception that ungrading affords acknowledgement, or validation, of their labor. A full definition of "labor," especially in the context of contingent faculty, is beyond the scope of this chapter. A glance at Inoue's *Labor-Based Contract Grading* underscores that literature on labor-based contract grading defines *labor* almost exclusively in terms of student experience, the "work done in and for a course" (pg. 76). In labor-based approaches, this work corresponds to the grade students receive. This focus leaves entirely unaddressed the work *teachers* do "in and for a course," and how conventional or alternative assessment practices amplify, or reduce, the value teachers' experience in their labor. Taylor's experience of finding ungrading the answer to a deep-seated desire for professional agency answers this concern, showing how the correspondence between value and affordances in ungrading may meet *teachers'* needs, as well as students'.

Taylor frames their working conditions as insufficiently compensated or incentivized within the department. Teaching four classes (often with 20+ students enrolled, for 80-100 students/semester) is typical in the UTK First-Year Composition Program for full-time NTT lecturers like Taylor. Yet the crux of the problem is the tension between compensated and uncompensated work, or between work is that is valued and unvalued. Taylor explains: "When you're a lecturer …. time is so precious because you're teaching so many classes and you're trying to do all this other stuff that [you're] not compensated for, which is different from being on a tenure track, where the research and writing that [you're] trying to do is not part of the teaching work that we are evaluated on or paid for." Here, the problem is not just multiple, overwhelming responsibilities. The problem is the evaluative and compensatory structure of NTT faculty, which both invites and discourages "research and writing" as part of faculty work. Ongoing scholarly work offers NTT faculty several benefits, such as potentially teaching more effectively and laying the groundwork for promotion or employment in a tenure-track position elsewhere (Colby and Colby, 2017, pg. 61-62). This work, however, is as Taylor observes uncompensated; all compensation is — ostensibly — for teaching, in contrast with research or service, as with a TT/T role.

The comment about time gestures towards the way that conventional grading assumes a value structure which undermines instructors' agency and control over their labor, while alternative assessment (potentially) restores agency, validating the labor of instructors. Taylor explains:

Table 5.9 Values, listed by prevalence — Taylor, 2020-2021

Values by prevalence	No. of references
Teacher agency	9
Fairness	7
Recognizes students' humanity	5
Empowering students	2
Comes alongside students	0
Rhetorical effectiveness	0
Strong pedagogy	0

Table 5.10 Perceived affordances of ungrading, listed by prevalence — Taylor, 2020-2021

Affordances, by prevalence	No. references
Instructor labor	13
Attends for social inequities outside of race	5
Attends for social inequities — COVID	4
Student labor	2
Interacts with pedagogy	1
Develops student-teacher relationship	0
Antiracism	0

> [When I was ungrading] I was able to actually focus my time and my comments on what I have always wanted to focus them on as a teacher, I felt like I was able to spend my time actually commenting on the work that [students] were doing as opposed to, you know, I have to justify a grade or I have to take five points off because [students] didn't do the blah blah blah which I don't even care about, like a Works Cited thing or whatever. And so I was able to actually focus on the content and the organization of how they were articulating that content, which those are my goals as a writing instructor, especially for students who are that young.

I think two things are going on in this passage. When Taylor notes that ungrading made it possible to "focus my time" on pedagogical tasks they believed were important, they're speaking literally. They have 100 papers to grade, and ungrading (re)distributes the hours and minutes in ways that are less burdensome. But Taylor is also referencing broadly the ability to do their work in the way that aligns with their goals and beliefs as an educator, the ability to teach with integrity. Taylor brings a given set of values and commitments into the work of writing education, and the kind of assessment used — conventional grading or ungrading — affects whether Taylor goes about their work in a way aligned with their values, or misaligned. Just as the goal with liberatory assessment practices such as labor-based grading aim to give students control over how they do the work, assessment affects whether educators have autonomy and control over their labor, a desire clearly articulated in Taylor's comments here.

In conventional grading, educators' labor is often *mis*aligned with their values. Notice Taylor's repeated use of "have to:" Taylor feels they "have to" justify the grade given to students, or they "have to" penalize students for a particular error. This phrasing frames grading as fulfilling obligations set by a higher authority, much the same way that educators in the 1930s and 1940s in the UTK English Department graded according to standards established by the larger institutional hierarchy. These standards sit uncomfortable with Taylor, who also describes being forced to penalize students for things they "don't even care about," such as a Works Cited[97]. Conventional grading, which emerges from and reifies the institutional architecture imposes on teachers a narrow set of values around what kind(s) of writing and learning tasks are important. Because grading participates in and reinforces hierarchical structures, the work of grading robs Taylor as an instructor of their own agency.

Ungrading, in contrast, is not keyed to same value structure(s) as conventional grading is and so Taylor experiences it as giving them "time" to do their work in the way they want to, or in other words, control over what work they do and the values they honor in that work. As Taylor explains, ungrading "focus[es their] time" in the way they wanted, on providing students

[97] I did not ask Taylor to specify what they meant by a "Works Cited thing." As an experienced educator myself, however, who personally has little patience for questions of formatting, I read this comment as referring not to the ethical practice of citing sources — certainly an important value to maintain in our assessment of students' work and feedback to them! — but to the much more finicky work of requiring particular documentation, such as MLA or APA, which may often exceed students' capacity as first-year writers and may, as Taylor implies, interfere with teaching other, more important, skills. I would also point out that Taylor is not the only instructor to focus on the curricular emphasis on citation practices as conflicting with the values/goals they bring to ungrading; as I discussed in the previous chapter, Finley likewise finds citation as a key curricular emphasis problematic.

constructive feedback and coaching them in higher-order thinking skills, in contrast with the lower-order thinking skills implied in correctly producing a "Works Cited thing." That Taylor describes this refocusing as something they've "always wanted" is key, as it stresses that Taylor's longstanding interest in reclaiming professional agency contributes to their use of ungrading, drawing out its potential to validate their labor. Elsewhere in our conversation, Taylor explains that because ungrading enables them to spend "much more time commenting on and evaluating student work in a way that matches up with my values as a person and my pedagogy as an instructor," it made them feel that they were "teaching in line with my values." Again, that Taylor has "time" to conduct their work in the way they want to conduct their work is evidence of ungrading's effectiveness at validating teacher labor, reworking the classroom space to open up new possibilities for instructors as well as students.

Ungrading's (re)structuring of labor to put control in teachers' own hands is what attracts Taylor to it long term. Taylor explains: "I loved ungrading. A lot. I don't think I'm ever going to go back to traditional grading. For me, it was a really good experience, but yeah, I've — it made me feel better about my teaching, it made me feel better about my time management, it made me feel better about the work that I am putting out into the world and serving other people, so yeah I loved it." What is striking about Taylor's response is its emotional warmth, suggesting greater attachment to or identification with professorial or pedagogical work than previously, within a conventionally graded approach. Taylor finds the value they place on agency acknowledged and rewarded within ungrading; ungrading makes it possible for them to teach authentically, in a way that carries with it the intrinsic reward of feeling good "about the work I am putting out into the world." Especially given NTT faculty's lack of compensation for their work and corresponding desire for agency and control, that ungrading, for Taylor, rewards instructors' labor as well as students' makes it a more lasting form of assessment[98].

Ungrading, then, depends on educators' values or commitments for achieving particular ends. As an alternative to conventional assessment, it *may* be equitable, more attentive to students' and teachers' labors and aligned with best practices in education, but it requires that teachers also care about equity and good teaching to reach those goals. Additionally, the outcomes or affordances of ungrading that stand out to teachers are often a reflection of their own commitments, suggesting that educators may find ungrading a more fluid, adaptable practice for reaching goals central to their professional identity than conventional assessment practices.

This brings me to the next section of this chapter, in which I address the experiences of two educators — Taylor and Sterling — who used ungrading for multiple semesters. Through this longitudinal study, I show how continued engagement with alternative assessment practices may alter or deepen instructors' commitments in ways that change the effects of ungrading in first-year composition classrooms.

[98] Of course, other NTT faculty may have a different experience with ungrading than Taylor does. As Taylor points out, while they valued ungrading in part because it gave them greater time and control over their work, lecturers with more experience at UTK — decades into their teaching career — might find the labor of getting started with ungrading, moving into a new assessment ecology, more trouble than it is worth.

Ungrading's affordances during 2021-2022

The data in this section was collected during AY 2021-2022, as part of the Equitable Assessment Pilot. Both Taylor and Sterling participated in the Equitable Assessment Pilot and subsequently agreed to an interview, so I was able to compare their experiences using ungrading over time. (I focus on Taylor and Sterling here for the sake of time, and to provide useful comparisons with the examples I used above; they are, of course, not the only teachers who have multiple years' experience using ungrading.) This section focuses on how Taylor's and Sterling's teacherly values, and the affordances they perceive in ungrading, change from the first year (2020-2021) to the second (2021-2022) and sheds light on how ongoing experience with ungrading affects teachers' perception of its outcomes.

I took several steps to analyze this change over time. First, I documented the change in frequencies of both values and affordances in coding, comparing the emphases which appeared in Taylor's and Sterling's conversations in 2020-2021 to the emphases in 2021-2022, using the charts in the section below. I also made qualitative comparisons among core codes, or those which remained emphasized, or changed in emphasis; for instance, because "instructor labor" is central to Taylor's experience in 2020-2021 but subsequently decreases, I listed key themes from one year to the next, that are reflected in the analysis below. In doing this, I relied on Saldana's guide to longitudinal coding and data summary, focusing my analysis on mapping out 1) increase and growth, 2) turning points, 3) decrease(s), and 4) missing or absence; though I didn't use these as "codes" per se, meaning I didn't enter them into NVIVO, I analyzed these shifts as I made qualitative comparisons among the existing codes. I also considered how these shifts may have occurred in the *external* or institutional context, identifying instances of increase, turning points, and decrease that affected Taylor's and Sterling's experience.

The most obvious turning point, of course, and the one with the largest impact on this study, is the existence of the Equitable Assessment Pilot — the one-year, department-funded support for first-year composition instructors using alternative forms of assessment — in 2021-2022. Though ungrading was informally supported in 2020-2021, the existence of the Pilot the subsequent year, along with funding for educators who participated in the pilot, served as a marked difference. I discuss how the Pilot functioned as permissory, and provided safety, for instructors in a way that allowed them to access both institutional and peer resources to craft a workable strategy in line with their values in the subsequent chapter, but here, I want to briefly note that for educators who participated in both 2020-2021 and 2021-2022, the pilot ensured space and opportunity to explore their values and affordances in fruitful ways. Sterling, remarking that "a lot of graduate students are handed particular rubrics or assessment tools that they have to use," expresses appreciation for the "institutional support" provided by the pilot:

> I think the reason that I decided to do the pilot again this semester was because I did want to change some things and I kind of wanted to make sure that I had a springboard and other people to talk to and resources to see um how those changes could work or give different ideas for changes to make and that sort of thing, so that was sort of why I decided to do it a second time, because I knew I was gonna -- I mean, not super significantly but significantly enough, change some things.

Taylor likewise describes participating in the Equitable Assessment Pilot again out of a desire to change some things:

> And I think like the reason that I decided to do the pilot again this semester was um because I did want to change some things and I kind of wanted to make sure that I had a spring board and other people to talk to and resources to see um how those changes could work or give different ideas for changes to make and that sort of thing, so that was sort of why I decided to do it a second time, because I knew I was gonna -- I mean, not super significantly but significantly enough, change some things.

What I want to stress is that, although this discussion describes the growth and development of Taylor's and Sterling's values around assessment and teaching, and the affordances they perceive in ungrading, they both attribute this growth and development to the institutional space opened up by the pilot — showcasing the importance of a hospitable architecture, as I discussed in the previous chapter, in sustaining the use of ungrading long enough for it to reach meaningful equity-related outcomes.

One final note. Since I only include Taylor and Sterling in this analysis, my longitudinal approach emphasizes teacherly trajectories; it is not a cross-section of data (Grosshoelme and Lipstein, 2016). What this means is that my emphasis in this section is not on mapping out distinct, generalizable differences between the crew who ungraded in in 2020-2021, and the crew who ungraded in 2021-2022, nor differences among people generally in their first and second year ungrading. I focus on describing how Taylor's and Sterling's experiences grow, change, and develop, which can speak to how the affordances of ungrading might also change and develop, depending on how long educators use ungrading and/or the institutional supports that are available to them as they do so. Because of this choice, the subsequent section focuses on first Sterling, then Taylor, walking through one instructor's experience at a time to more fully explore the individual, idiosyncratic process of change.

Sterling (2020-2021)

As Table 5.11 and Table 5.12 demonstrate, the values and corresponding affordances that characterized Sterling's first year using ungrading — attention to antiracism, grounded in a commitment to empowering students; and appreciation for ungrading's attention to instructor labor — shifted, the first deepening and increasing, while the second decreased. To demonstrate this shift, I have bolded the items which show a notable increase (for instance, empowering students expands from 21 references to 32), and bolded and italicized the items which show a notable decrease (instructor labor drops from 20 to 13 references). (*Notable* refers to a shift of 5 references more or less, between one year and the other.) Please note that though the tables are organized by frequency, I have emphasized the second year (2021-2022) over the first (2020-2021), to better emphasize where Sterling — and later in this chapter, Taylor — landed as a teacher, rather than where they began.

Similarly, this analysis in Table 5.12 shows that as Sterling's attention to ungrading as relieving instructor labor dwindles, their sense of ungrading as cultivating antiracist action expands and develops,

Table 5.11 Change in teacherly values, listed by frequency — Sterling, 2020-2021 to 2021-2022

Value, by frequency (Sterling)	No. References (2020-2021)	No. References (2021-2022)
Empowering students	**21**	**32**
Teacher agency	15	18
Recognizes students' subjectivity	6	10
Strong pedagogy	7	10
Fairness	3	4
Rhetorical effectiveness	***9***	***3***
Comes alongside students	1	3
Self-care	3	1

Table 5.12 Change in perceived affordances of ungrading, listed by frequency — Sterling, 2020-2021 to 2021-2022

Affordance, by frequency (Sterling	No. References (2020-2021)	No. References (2021-2022)
Interacts with pedagogy	***26***	***18***
Antiracism	**5**	**18**
Instructor labor	***20***	***13***
Student labor	3	6
Attends for social inequities — COVID	5	4
Attends for social inequities outside of race	0	2
Develops student-teacher relationship	0	2

demonstrating how the affordances linked to ungrading may expand and grow the longer instructors use ungrading.

One important turning point for Sterling, in addition to the Equitable Assessment Pilot, is that during 2020-2021, they taught a first-year composition course which enrolled more students who did not use standardized American English as their home or primary language. A turn which resonates with Sterling's interest during 2020-2021 in students' writerly identities, the new opportunity opens space for Sterling to articulate more fully the role of ungrading in inviting critical conversations about how language is judged and encouraging students' use of diverse Englishes in first-year composition. This commitment comes to characterize Sterling's pedagogical framework overall. Describing their experience in 2021-2022, Sterling stresses the importance, in using ungrading, of purposeful attention to antiracism:

> One thing for me is that I think that it's irresponsible and becoming impossible to not think critically about assessment. I think we are in this moment where we cannot just continue to use assessment in the way that is reproducing very negative ideas about literacy and language and identity and really we need to use assessment as a way for equity and inclusion in the writing classroom.

Ungrading, and specifically direct conversations about antiracism, is for Sterling a way of enacting their commitment to antiracism. As the conversation goes on, they describe their deliberate efforts to ensure this inclusion through in-class conversations. Noting that their section of first-year writing "is centered around difference in students," Sterling argues that "it became much more important to say something about that [difference], where I couldn't just ignore it like I had in the past." This comment picks up on observations Sterling made at the end of their first year ungrading, about the need to more directly address with students how their language is judged across multiple contexts and engage them in conversations about power, language, and evaluation. What is significant here is that in their second year ungrading, Sterling doesn't only plan for these conversations but starts having them, intentionally building space in their curriculum to invite students to think about "equity and language ideology and how that's always functioning in the language classroom." Sterling explains that their goal in these conversations is to "get students thinking beyond my class, to the ways their writing is being judged in other classrooms as well" — a vision or goal which echoes similar goals from the first year, but which Sterling clearly perceives themselves as reaching more fully or completely, in the context of this new class. Building on their prior commitment to empowering students, these conversations demonstrate a commitment to praxis, deploying ungrading, and the various conversations it sparks, to engage students in building awareness of the judgments around language and enable them to make savvy choices across contexts.

Sterling relies on ungrading not only to open up conversations about power and language but invites students to act on what they learn. Sterling explains,

> So for me, contract grading [fit] with [my] course design because something I was trying to do was encourage people to explore English, and explore literacy, in a fuller range than just standard English or standard academic English. So contract grading actually made

> that more possible, because students felt more comfortable taking those kinds of risks, and writing more in a way more to explore identity and develop as writers rather than just writing for a particular grade.

Echoing themes from Sterling's first year using ungrading, especially the attention to how students develop as writers and their willingness to take risks, this passage underscores the particularity of Sterling's commitment to empowering students through the use of ungrading as an antiracist tool. The fact that ungrading does not penalize students from deviating from a perceived or Platonic ideal means, as Sterling explains, that students feel comfortable trying out ways to write other than "standard English or standard American English;" a "fuller range" of possibilities opens up in language use, especially in academic settings. But significantly, this passage captures benefits not only for Sterling's multilingual students but also for the wide range of students, anybody enrolled in the course who uses something other than standardized American English.

Another change between 2020-2021 is that, compared with the first year, Sterling puts less emphasis on ungrading as validating their own labor, or shaping their identity as an instructor. Instead, they frame ungrading as a way to express and affirm their professional identity and values. I speculate that to some degree, this shift is explained by the reality that, as a tool to sustain and reward instructor labor, ungrading has met its purpose. Having opened up space for Sterling to explore their values, ungrading is now something Sterling feels more confident about, thinking beyond their own identity to wider professional outcomes of any given assessment choice. Sterling says:

> And now everything feels so resonant to me, I — and it because of the pilot, and because of all the support from the writing program, I do not feel vulnerable using this type of assessment, because I feel like I have that freedom to experiment and to really take a stand with the assessment that I use.

I find Sterling's attention to the "resonance" between their use of ungrading and their pedagogical commitments telling, especially in light of their sense that they can "take a stand" by using ungrading. A deeply value-laden phrase, *taking a stand* gestures towards the way that assessment has become for Sterling an outworking of established commitments within the framework of their teaching practice; it denotes stability and endurance over the long term. On the one hand, this commitment affirms that, even as Sterling continues to rely on ungrading in the classroom, it rewards and validates their labor, aligning it with their teacherly identity in satisfying and affirming ways. On the other, the language of stability is distinct from Sterling's framing in their first year of grading. Then, they stressed ungrading as a more dynamic, formative process, arguing that they found themselves as an educator. Now, having used ungrading for a longer period of time, they frame ungrading as solid or stable, a rooted ground from which to articulate and practice their pedagogical commitments.

In coming back to ungrading for a second time, then, Sterling builds on their prior commitments. Where, in the first year, Sterling found ungrading a space to construct their teacherly identity and pinpoint equity-related goals, here, they carry out those goals, using ungrading to invite

conversations about language and power into the classroom, more fully meeting the needs of students enrolled in their course. In the previous chapter, I addressed how the Equitable Assessment Pilot gave instructors the sense of permission or freedom to take up ungrading; what Sterling's experience here suggests is that instructors make productive use of that freedom to explore and strengthen their own values, reinforcing both the value of developing a hospitable programmatic architecture and the importance of sticking with ungrading over multiple terms.

Taylor (2021-2022)

Sterling experiences ungrading as deepening and (re)arranging prior commitments but, crucially, Taylor experiences longer use of ungrading as drawing out new, or at least hidden, commitments. As Table 5.13 shows, though Taylor retains key values such as instructor grading, new ones emerge. For Taylor, the pilot, and the extended use of ungrading, draws out a commitment to education that recognizes students' subjectivity and individuality, and the corresponding use of ungrading to meet social inequities, in particular disability and difference.

While the perceived affordances Taylor finds, depicted in Table 5.14, in ungrading do not map perfectly onto their values, notice that especially in the second year, they value ungrading's ability to attend for social inequities outside of race and to cultivate a stronger pedagogy — both outcomes that reflect Taylor's keen awareness of the ways that students come into their class as different people, with different backgrounds, and Taylor's desire to meet those students' needs in their teaching.

I want to call attention to two key differences in Taylor's second year (2021-2022) ungrading, compared with their first (2020-2021). The first difference is Taylor's new attention to ungrading as recognizing student subjectivity — or for Taylor, ungrading as an affordance for universal design. I developed the code "recognizing student subjectivity" to capture teachers' valuing of students' complex life experiences amid the pandemic, but for Taylor, this value focuses particularly on ensuring access for neurodivergent and disabled students. Speaking of "students [who have] accommodations from the Student Disability Services," Taylor frames ungrading as ensuring those students are "just as capable of getting the A in the course as a student who doesn't have any disabilities or accommodations." For Taylor, ungrading works as "universal design, a universal[ly] designed grading system that works as well as it possibly can for neurodiverse students." Reflecting Inoue's interest in ensuring that the "full range of grades" (3) is equally available to all students, Taylor's values here extend that to disabled and/or neurodivergent students, grounding educational choices in the desire to see all students succeed. If Taylor is centrally interested in meeting students where they are, ungrading becomes or functions as a tool that makes it possible for them to do so.

I asked Taylor to follow up about their perception of ungrading as universal design. Taylor replied:

> It feels to me like it is a design — a grading design that has as few barriers to access as possible. So that it isn't something that, you know, like a handicap bathroom is specifically made for like a wheelchair user or that's specifically made for somebody with a particular type of disability, but that it says like hey, how do we create a system

Table 5.13 Change in teacherly values, listed by frequency, Taylor, 2020-2021 to 2021-2022

Value, by frequency	No. references (2020-2021)	No. references (2021-2022)
Recognizes students' subjectivity	**5**	**15**
Teacher agency	9	8
Fairness	7	8
Strong pedagogy	0	3
Empowering students	2	1
Self-care	1	1
Comes alongside students	0	0
Rhetorical effectiveness	0	0

Table 5.14 Change in perceived affordances of ungrading, listed by frequency, Taylor, 2020-2021 to 2021-2022

Affordance, by frequency	No. references (2020-2021)	No. references (2021-2022)
Interacts with pedagogy	**1**	**9**
Attends for social inequities outside of race	5	8
Instructor labor	***13***	***7***
Student labor	2	6
Attends for social inequities — COVID	4	4
Develops student-teacher relationship	0	4
Antiracism	0	1

> that has — that considers everyone or as many different types of people as we can possibly consider, and make it work for — if not all, then as many as we possibly can make it work for in a system?

Taylor here recognizes that some students may experience "barriers to accessing" or reaching particular grades, on the basis of their dis/abilities, a profoundly individual, subjective approach that acknowledges students come from different places. At the same time, Taylor is not interested in "fitting" students[99] for a presumptively neutral grading approach but in fitting the grading approach to all students. This ensures that wherever students are located, whatever places within the system they inhabit, they are equally able to access success, a framing in line with disabilities studies research on universal design and accommodations. Reflecting on student experiences, Taylor suggests:

> If a student has a lot of anxiety around grades, it's helpful for them, if a student is really an overachiever and really really wants to get that one hundred percent A, it's very achievable for them, if a student is someone who works really hard and is not an awesome writer and that's something that they really struggle with they have an equal opportunity, so there's very few barriers to be able to access that A.

Because Taylor's approach to ungrading, like many (though not all) at UTK hinged on whether students "did the thing," measuring success by labor or engagement in a course (c.f. Carillo 2021), students could redo assignments they missed — which, though time-consuming, allowed students to overcome prior errors, regardless of whether those errors stemmed from students' fluency in writing, their mental well-being, or even their inattentiveness[100]. Taylor further explains that they also provided students multiple ways to complete assignments — in particular, different ways to complete a final assignment (outside-of-class engagement with the course theme) — that allowed for student choice among a wider number of alternatives, depending on students' interest and dis/abilities.

That ungrading, for Taylor, affords greater access to the course may come as a surprise. Recent research has suggested that ungrading may actually be *less* accessible, as it requires students to adjust to a new activity system (Kryger and Zimmerman 2020) and assumes labor as an equally available resource (Carillo 2021). While I don't have the space to address the dispute here, I do want to make two observations. First, my research focuses on *teacher* experience, not *student* experience; it is possible that despite Taylor's emphasis on ungrading as universal design, students did not encounter ungrading as enabling greater access to the course. More research in this area is needed.

---

[99] I drew on Bennett 2022 (ENGL 295 class presentation) for the language of fit/misfit.

[100] At the same time, framing ungrading as a way to overcome certain errors — for instance, perhaps poor or incomplete summary of secondary sources, something Taylor mentions — falls back into quality-based approaches to assessing student work, even though no grade is given (Inoue 2014). I mention this not to critique Taylor but to point out that the road to ungrading is complex and long, requiring ongoing engagement with and interrogation of teachers' values in writing.

Second, much of the criticism in both texts centers the labor logs — briefly, a record kept by students of the quantitative amounts of time spent on various tasks for the course (Inoue, 2019, p. 105) — as a chief example of contract grading's inaccessibility. Taylor, like other educators at the University of Tennessee, does not use labor logs and instead focuses on whether students "do the thing," completing to pass/no pass specifications the core assignments in a given course[101]. Additionally, like many other instructors at UTK, Taylor has a default A/A-, so that students are not asked to do "extra work" to move up from a B, as in Inoue's approach and other popular forms of contract grading. This approach gets around, at least in part, concerns about the labor logs as quantifying and even calling attention to the ways that some students need more time, or different flows of time, to complete their work . The approach may also highlight how, especially for teachers who are, like Taylor, invested in accommodating students who are disabled and/or neurodivergent, ungrading crafted with attention to universal design may actually be — though is not necessarily — accommodating.

Just as Sterling's emphasis on ungrading as an antiracist tool to achieve their goal of empowering students spills over into their wider pedagogy, so Taylor's emphasis on ungrading as universal design comes to infect their whole view of the course. Reflecting on two students, who came to the course with vastly different motivations and skill levels that affected their writing ability, Taylor notes that with ungrading, "you either did it or you didn't do it, and if you didn't do it then you do it again — that made both of them like actually engage with the assignment in the ways that they needed to. So for me that was an amazing example of how it was working for different — not just for one type of student, you know?" Unsurprisingly, given the nature of universal design, the assessment structure(s) that correspond with Taylor's more pointed interest in supporting students who enter the course with accommodations also support students who may not have or need those accommodations, students who for mundane reasons – – hypothetically, a lack of preparedness for the writing course, a big chemistry test the week before a paper is due, a computer that breaks unexpectedly. Taylor's commitment to supporting students' subjectivity, acknowledging there is not just "one type of student" in their course, comes to drive their use of ungrading across the whole course, so that it functions to support and refine their pedagogical practices more holistically.

The second observation I want to make, briefly, is that as with Sterling, Taylor's interest in ungrading as a way to validate their own labor, decreases. This value is not erased, though. Taylor notes:

> [With ungrading] it's definitely the work load go[es] down, and I think maybe more importantly is I feel like I'm — like I feel better about how I'm spending my grading time, so even if it took the same amount of time like I think it still matters that when I'm ungrading and giving this feedback in a specific way, I know that I'm giving my students feedback that will help them understand and learn from what they're actually writing.

[101] This is an example of the way my own ungrading materials — a quasi boss text, organizing educators' decisions about ungrading, as I note in the next chapter — affect educators' experiences with ungrading. I have never used labor logs; when I first switched to grading contracts, in 2019, I perceived them as complicated busy work. Because my approach served as the template for many, though not all, educators, nobody else at UTK that I know of uses labor logs either.

> And not — and that makes me feel like that is time well spent rather than oh, I'm just grading to give them this number that the university makes me give them, you know, which is kind of how I approach it.

On one level, the affordance or opportunities ungrading opens up for Taylor are, quite simply, a more manageable workload, ensuring that their time (as Taylor notes earlier in this chapter) is "well spent." But on a deeper level, Taylor returns here to themes from 2020-2021, describing how they "feel better about" their work as an educator, fulfilling their goals rather than performing undesired work, or adhering to work processes and institutional expectations, that are misaligned with their goals. That Taylor would at once continue to feel good (better) about their work while ungrading, even as they deepen and turn their attention to other goals, more student-centered than their own teacherly agency, is unsurprising, given that my argument in this chapter is that ungrading enables teachers to achieve their values and ways of being important to them in the classroom. The more Taylor (and Sterling) use ungrading, the more clearly they perceive their pedagogical goals, and they turn their attention in that direction. Yet because their practice is, increasingly, aligned with their values rather than institutional performances, they naturally continue to feel good about their work, to maintain their integrity, rather than teaching in a way out of sync with the larger institutional expectations.

Taylor, even more than Sterling, initially turns to ungrading because it meets their goals of greater autonomy in the classroom. But Taylor too finds that a second experience with ungrading deepens and complicates their sense of what ungrading is *for* — its potential to work for equity, in this case, dis/ability justice and accessibility. Together, the two teachers' experiences suggest that the longer instructors use ungrading, the more clearly they see its potential and put it to good work in writing classes. This outcome highlights the importance of deliberate, ongoing institutional engagement with antiracist and assessment theory. Additionally, the importance of using ungrading over a period of several semesters affirms the value of direct institutional support — the hospitable architecture discussed in the previous chapter, as that structure enables teachers to take the risks necessary to continue to invest in the practice for the time necessary for it to bear fruit.

Conclusion

Most research on ungrading has framed its affordances as (fairly) straightforward. Ungrading produces the conditions for good learning, greater risk-taking, and more creativity among students; ungrading cultivates an equitable space where conversations about language and power are possible. Valuable as this research is, the data I review in this chapter complicates it. For instructors at the University of Tennessee, as the program gradually turned towards alternative assessment, the affordances of ungrading varied, depending on their social and institutional position(s) and their deep-rooted beliefs about writing and (writing) education. UTK instructors at first took up ungrading out of a desire to respond to the exigences of COVID — the urgency of accommodating, in some way, the disruption students and teachers alike experienced during Summer and Fall 2020. Yet this response was influenced by deeper assumptions about what writing education should *be* — accessible, equitable, and centered around students' subjective needs. As the first months of COVID receded, this connection clarified: the affordances UTK instructors see in ungrading correspond with their own values and commitments. Instructors who

value autonomy in their teaching see ungrading as validating their labor; instructors who value equity, as inviting justice-oriented conversations.

Additionally, the longer instructors use ungrading, the more complex the relationship between their deep-seated values and ungrading becomes, prompting instructors to lean into the potential ungrading has to cultivate more just, liberatory spaces for assessment. I suspect that more time and space — especially with the sort of direct, purposeful departmental support that the Equitable Assessment Pilot offered — offers instructors the chance to reflect more deeply on their pedagogical commitments, try out new changes, and see growth. Ultimately, the give-and-take between teachers' values and the affordances of ungrading gestures towards the ecological nature of assessment. Educators are "enmeshed" (Byrd, 2019) in a web of values, institutional and social obligations, and assessment practices, shaping who they want to be and how they want to act in the classroom; this web in turn nourishes and gives rise to particular formations or experiences of ungrading, different journeys, with different ends.

Conclusion

I have, in this project, told the story of how grades and grading emerged in the First-Year Composition Program at the University of Tennessee, Knoxville — and then, nearly a hundred years later amid the overlapping social crises of the COVID-19 pandemic and the resurgence of Black Lives Matter, how the program facilitated contingent faculty's choice to use more equitable forms of assessment in their composition courses. Along the way, I use this case study to explore the role of institutional architecture in shaping, and supporting, both grading and ungrading in first-year composition courses, and the influence of individual teachers, working within those institutions, on the direction and impact of ungrading. At the University of Tennessee, I show how grading, in the early years of First-Year Composition, bowed to institution mission and pressure, serving primarily to measure the perceived quality of students' work, using a Platonic ideal of American English as a standard, in ways that effectively barred minoritized students from (succeeding at) college and constructed the students who remained in the classroom as deficient writers. A century on, however, the architecture of the First-Year Composition Program proved hospitable to ungrading, relying on alternative boss texts such as ungrading models, strong peer and hierarchical relationships, and ample personnel and textual resources to invite ungrading. As first-year writing educators turned towards ungrading, the values and commitments they brought to the practice — whether a desire for greater fairness, for equity and antiracism in teaching, or for their own increased autonomy — surfaced through their use of the practice. At the same time, teachers who used ungrading for multiple semesters found that their commitments and values deepened and changed, the longer they used the practice.

Threaded throughout this history is the following key question: Where did grades and grading come from historically in the University of Tennessee First-Year Composition Program, and how can the institution provide meaningful support for alternative, equitable forms of writing assessment, especially among the contingent faculty who teach these courses? When contingent faculty come to use alternative assessment, how do such practices have an impact in their classroom? This focus allows me to grapple with questions about the role and importance of institutional support in maintaining both conventional and alternative forms of assessment, and the impact that individual teachers' values and choices have on assessment practices, contributing to their success in cultivating more learning-centered, equitable environments. In response to these questions, I find that at UTK, that grades/grading originate from institutional needs, not students' needs; and that accordingly, as the institution prioritizes developing programmatic architectures hospitable to ungrading, alternative assessment practices are more likely to flourish among contingent faculty. I also find that ungrading, rooted in faculty's values and beliefs about teaching, has fluid, variable outcomes, from supporting learning to empowering students, and has stronger impacts the longer it is used.

These findings mean that institutional commitments, and individual teachers' commitments and values, as articulated within particular institutional architecture(s), are key in determining whether ungrading takes off, and what outcomes it leads to. Grading is not — as we might like to imagine — an objective articulation of authoritative judgment on students' work but a deeply subjective process influenced not only by teacherly values but by institutional missions and commitments. Changing grading — moving towards antiracist, liberatory alternatives —

likewise requires simultaneous attention both to institutional mission and to our own values and commitments in the classroom. Ungrading is not (only) a matter of individual teachers' choices about their individual classrooms; ungrading is a deeply institutional process, invited into, complicated by, and/or at times barred from, the spaces that our writing programs, departments, universities, and local and state communities construct around the work we do within institutions. Ungrading is not a plug-and-play model, with clearly-defined and predictable results — among them, a better and more racially equitable learning environment. Ungrading is instead a deeply variable practice, capable of working towards equity but also powerfully (re)shaped by the goals that teachers bring to the work. As writing teachers, administrators, and scholars invested in more equitable approaches to assessment, then, this work invites us to think critically about the values we bring to the work of ungrading, and the institutional spaces we inhabit — or leave behind — to do the work most effectively.

These findings, of course, emerge out of a case study, a method known, as I discussed in Chapter 1, for its limited ability to speak more generally, to contexts beyond its own. The work I conducted in this project is deeply situated within particular contexts not only institutionally — an R1, state flagship in a Southeast Appalachian city with ongoing racial violence, including police violence — but also socially and historically. The bulk of both data collection and analysis occurred during a period of heightened awareness of injustice — a period which has, perhaps unsurprisingly, been followed by a strong conservative pushback, throwing cold water on efforts among university faculty, graduate students, and staff to engage in antiracist — or even learning-centered — practices. As I write this conclusion, states such as Florida, Texas, and North Carolina are increasing surveillance of universities — for instance by threatening tenure and employment, defunding diversity and equity initiatives, or ramping up state-level review of faculty curriucula, with governors' offices requesting information about which courses assign or teach work perceived as critical race theory. Even institutions are more reluctant to take up, or continue, equitable approaches to teaching, research, and service, often opting out of processes considered to be more equitable, such as test-optional admissions. The University of Tennessee, for instance, piloted test-optional admissions through AY 2020-2021 and 2021-2022, then abruptly rescinded the pilot beginning in Fall 2020 ("Test Optional Pilot Program Continuing," 2021, "Test Score Policy"). We cannot recreate the openness to change made possible by the pandemic, outside of those particular questions; and as the years pass, we may find that the changes outlined here — the turn towards a more hospitable programmatic architecture, or the willingness to reflect on and change one's own beliefs about teaching — no longer meet with a receptive audience. Individuals, academic institutions, and social contexts may all prove more resistant. Because this project emerges from a unique turning point in our civic experience, its ability to intervene in the teaching and administration of other universities, especially at the longer term, is limited.

At least in the short term, however, my hope is that this work matters, along three dimensions: a call to reckon with the history of and function of grades/grading in first-year composition courses, to situate our assessment choices within the institutional structure(s) that shape them, and commit to making whatever changes we can, in pursuit of more liberatory approaches to writing education and assessment. First, my work invites us to be honest about the origins of our teaching and assessment practices. Grading is not (as we might hope) a way to uphold

excellence, maintain high standards, or push our students to succeed. It is not even a viable way to track students' work within an institution, given that (as my project has demonstrated) it is a highly arbitrary practice subject to the changing mission and pressures of the university spaces we inhabit. What history shows is that grading is an exertion of institutional power and supremacy, a way to rank students — to rank *people* — by their perceived intelligence and capacity, determining who succeeds and who does not. Grading perpetuates and extends the disciplinary function of the university, enculturating students into its own norms. This reality does not mean that all teachers who use grades and grading should instantly stop. As I hope my project has made clear, instructors' motivations and goals for ungrading both vary widely and shape the outcome, making a quick dash away from grades and grading, without any coherent idea where we are going and why, unlikely to bear much fruit. But if we do still use grading, whether in whole or in part (and to be clear, even those of us who use *un*grading still grade, so long as we are asked to attach a number to students' work at the end of term) acknowledging the historical realities of that practice — its rootedness in discrimination and institutional prestige — is key to making choices about who we are in our professional lives and how we interact with students, colleagues, administrators, and the institution that are ethical and judicious.

This point brings me to the second reason I believe my project matters: its invitation to (re)consider our obligation towards, and the possibilities within, the institutional spaces that we inhabit, in our work as writing teachers and administrators. I want to be up front that ungrading, like so many of the other institutionally-situated reforms we practice is a partial solution, not a complete one. I noted at the outset of this project that ungrading is partially, not wholly, antiracist, in that it withdraws penalties for linguistic/rhetorical choices that fall outside of white, standardized forms of academic English without (necessarily) affirming the power of choices grounded in students' cultural traditions. Ungrading is also used within existing systems; it does not recreate them and so is limited in its ability to transform them. Put another way, if ungrading offers a way to renovate the academic spaces we inhabit so they are more liveable, for students and teachers alike, that renovation is always, of necessity, incomplete. Most universities in the United States are founded on a history of exclusion, domination, classism, racism, and sexism and other forms of discrimination. Unless that foundation is razed and the structure rebuilt, such institutions will continue to perpetuate injustice and harm. Rooted in institutional contexts, this project lacks the scope to explore what wholly just forms of assessment could look like and the role of assessment in (re)constructing educational spaces altogether new, or external to, current institutional systems. This work is important, however, and I hope that future research on equitable assessment will take it up more fully.

My aim here is more pragmatic, focusing on the benefits of reforms and renovations we can make *now.* So long as we inhabit existing institutions, with their discriminatory histories, how does ungrading reconstruct those spaces in ways that reduce the impact of such discrimination on students and teachers and amplify the benefits of education instead? How can ungrading make assessment, if not altogether liberatory, at the very least less harmful and exclusionary for raciolinguistically diverse students and their peers? Institutions are often key in sustaining education, given that their structure ensures we can work in coordination and more efficiently than on our own. In highlighting the pressure exerted by institutional mission historically, and the importance of hostile and hospitable architecture, I see my project as asking us teachers and

scholars to think creatively about how we can take advantage of the places we inhabit in institutions to sustain ungrading and antiracist assessment — identifying places of alignment between the goal of supporting ungrading and institutional mission; leaning on peer and hierarchical relationships to ensure the work of ungrading gets done, whether with formal or informal institutional support; and even considering *where,* institutionally, we can do our best work. All the good work done in university spaces is inherently institutional work; it happens through institutional channels, whether using or exploiting them, or it does not happen at all. The work of ungrading is not, fundamentally, a solely classroom-level decision, something for individual teachers to opt into or out of; the work is institutional — not only at the university level but also at the level of disciplinary organizations (among them Cs and CWPA), journals and editorial boards, accrediting agencies, and all the other institutions we cycle through during our career. I see this project as inviting us as professionals to think *institutionally* about that work, to lean into the structured, formal relationships within our professional homes and make something of those to support ungrading — and more broadly, alternative forms of assessment.

Because, ultimately, this is the third and most foundational reason that this project matters, the insistence that *It doesn't have to be this way.* We don't have to teach and assess students using methods that have their rootedness in white supremacy and systems of exclusion. We may be many years away from dismantling institutional systems altogether, but we don't have to accept institutional ways of being and doing things, simply because they come from the institution (though we may have to be patient and savvy about working for institutional change). We may, as I noted earlier, be heading into a period of greater resistance to antiracist or equity work, and greater state surveillance and authoritarian control, but we do not have to passively accept it. Which practice, exactly, we do use to assess students (or whether we assess students at all) matters, but less so than the deeper questions about the value(s) and beliefs that drive our work – – our beliefs about writing, writing education, and our own students. What matters is that we approach the work of ungrading — and teaching more broadly — with a commitment to supporting students' exploration of how their identity influences the linguistic choices available to them, and to building with them a relationship grounded not in judgment and control but in trust and respect.

My hope is that this project may serve as a map for readers interested in charting out a different path, one that pursues ways of assessing that acknowledge students' diversities, including linguistic, rhetorical, and/or cultural diversities; and ways of leading or organizing writing programs that protect and invite equitable assessment. Documenting the turn from rigorous, conventional grading practices towards ungrading, I locate the grassroots movement towards alternative, liberatory forms of assessment in the (largely) hospitable programmatic architecture of the UTK First-Year Composition Program and in educators' own commitments to student-centered, equitable education. Our experience at UTK, of course, cannot serve as a perfect map; each program or institutional location will have its own unique, geographic differences. Other universities may be better positioned to lean into the antiracist potential of ungrading, drawing out assessment strategies that validate rhetorical choices that speak to audiences and purposes beyond white academics and academia. Even so, ungrading is, as I noted above, a stop-gap measure; it cannot cure the far larger, deeply rooted ills of the university system or even first-year composition at large. Yet my hope is that by charting for readers the institutional and

individual processes in developing new forms of assessment, the project invites us towards practices that work for educators as well as students, that invite us toward an ongoing commitment to renovating and remarking our classrooms and programs in creative, more equitable ways. My hope is that this work makes it possible for you, the reader, to insist — in the institutional space(s) you inhabit, in the classes you teach and the relationships you construct with colleagues and with your students — that *It doesn't have to be this way.*

Towards this end, I leave this project by addressing three groups of readers — writing educators, administrators, and scholars — and mapping out the possible implications or next steps of this work.

Writing educators

For writing teachers, this work invites us not only to take up ungrading but to grapple with our values and goals in doing so. Ungrading, as I discussed in Chapter 5, is a fluid practice, taking different forms that reflect and bear witness to our motivations in taking up the practice. An ungrading practice grounded in authoritarianism, subjective judgments, and pursuit of rigor and elitism will not bear the fruit of ungrading at all, while even a hybrid or partial form of ungrading, oriented around a desire to construct a non-hierarchical, learning-centered, liberatory classroom space, may be more fruitful in reaching those goals. To the educators reading this project, I invite you to consider your own goals with ungrading: Where are you trying to go with the practice? What kind of classroom, or relationship with your students, do you want to create? Do you aim at cultivating a better learning environment? A more equitable one? Do you want to reduce your own workload? How satisfied are you with these goals? I show in this project that the outcomes of ungrading are closely linked with teachers' values and commitments; by reflecting on our goals with ungrading, we educators increase the likelihood that ungrading will bear fruit in our classroom.

As part of this work, I encourage educators looking to take up, or extend and expand, their use of alternative forms of assessment to identify practices that make it possible not only to do ungrading but think critically about the values and goals underpinning that practice. Educators may benefit from connecting with colleagues (locally or at other institutions), so together they can grapple with these questions, share notes and models, try out new ungrading models and report back on how they worked. Educators may also benefit from asking students directly about their experiences taking an ungraded course, thinking about whether students' feedback on ungrading lines up with educators' goals for ungrading. New educators may want to take advantage of the wealth of online resources for ungrading, such as Asao's (Labor-Based Grading Resources) or mine (currently hosted on the document-sharing site Craft), while more experienced educators may benefit, or support others in this work, by organizing and sharing their own resources — work that invites reflection on the beliefs and goals that ground our work. Educators looking to use ungrading for antiracist ends may benefit from pursuing the work, largely left undone in this project — locating, studying, and creating models of assessment that validate diverse linguistic or rhetorical choices. Ungrading can have equitable outcomes, if teachers bring the goal of working towards equity to the table; but equity work requires not only a passive withdrawl of penalizing perceived errors but active, affirming engagement of students' own voices and home languages. Alongside the labor of reflecting on teacherly goals, and

locating/sharing models for more pedagogically-sound, antiracist approaches to assessment, I urge the writing teachers who read this project, and who work to hone their ungrading practice into a fruitful, liberatory assessment approach, to be patient. Ungrading is not a practice realized all at once; rooted in our values, it bears fruit over time, as we come to understand and complicate our own values more fully.

Writing program administrators

For writing administrators, this research project serves as a reminder that interest (your own and/or other faculty members' interest) is not sufficient to sustain or deepen the use of ungrading in the program. A hospitable architecture — one that through boss texts, peer and hierarchical relationships, and work processes, among other structural features — is essential. A more hostile architecture may discourage or outright bar the use of ungrading among contingent faculty who, as Carter puts it, "work under people and within the system" in ways that, by limiting professional agency and making conventional grading — however inequitable — feel obvious. Developing systemic supports for or openness to more liberatory forms of assessment invites such faculty towards the risky, creative work of ungrading. Writing administrators, then, are tasked with the work of reflecting not only on their own practice but also on the practice(s) and structures of the programs and institutions they lead — a particularly challenging task, if they have been at the helm of those institutions a long time, or had a hand in their construction. Because institutions are constituted by our own ongoing activities as administrators who lead them, or educators who inhabit them, we come to identify with the institutional spaces where we dwell. This is doubly true for administrators, who exert more control, or at least more *visible* control, and may feel more attachment to those spaces.

Yet for administrators looking to cultivate a more hospitable architecture, my project suggests several possible questions to consider how a program may be hostile towards ungrading — and how it may be *re*made, to be more hospitable. What role do grades and grading play in a program? What institutional purposes do they serve? Do they support a particular narrative about the institution? Do they function as a way to collect data about the writing program and its impact on students? Do grades or grading mechanisms (in keeping with the architectural metaphors used to describe programmatic structures) function as a load-bearing wall within the writing program? Administrators are also invited to ask questions about how ungrading may emerge in the writing program — what work processes support (or discourage) it, what resources (including both social and financial) are available for instructors looking to take up and experiment with ungrading practices, and what strategies may serve to remake the inherent hierarchies of the university in ways that invite, rather than discourage, new and risky forms of assessment. Administrators hoping to support antiracist forms of assessment specifically may want to consider what programmatic resources or supports faculty need, in order to critically engage diverse linguistic or rhetorical choices. As I discussed in the Introduction, meaningful antiracist work requires engagement with, and affirmation of, the rhetorical power of diverse languages/rhetorics/cultural expressions. Some educators may need hands-on training or resources to feel confident in doing this work, while other educators, especially educators from minoritized groups, may already be equipped to do this work but need to know that the Writing Program has their back. Finally, writing program administrators should consider how they can support not only individual but *programmatic* assessment alternatives. It does no good to foster

more equitable forms of assessment in classroom spaces, only to reimpose discriminatory, authoritarian assessment norms when we evaluate the work students do in our program, and how well the program helps them do that.

The work of developing a more hospitable programmatic architecture is difficult. Like old houses, constructed and reconstructed over time, writing programs may aim to take up ungrading, only to find that prior decisions about curricula, programmatic assessment, professional development or teacher support, and financial investments cultivate an unintentionally hostile environment. Engaging questions such as those listed above — about the way institutional activities, curricula, and other existing commitments assume grades/grading as the default, with little or no recognition of ungrading; about the training and support that is, or could be, offered for ungrading; about the financial commitments the program makes both to grades/grading and to ungrading; about the (tenured/tenure-track) faculty who support alternative assessment — may help administrators take a proactive role in the turn towards ungrading. Much of what I learned about hostile and hospitable architecture, crucially, I learned by talking with contingent faculty, and I urge fellow administrators to reach out to educators at their own institutions, for context-specific information about possible points of resistance. The more questions administrators ask about the role of grades and grading in their program, and about how the institutions we inhabit are arranged to invite contingent faculty to take up ungrading, the easier it is to partner with colleagues in cultivating a programmatic architecture hospitable to liberatory assessment.

Administrators, especially those established in the field, have the additional responsibility of thinking institutionally, not only within but outside their writing program, as I discussed above. This project invites us to (re)consider how the work of writing assessment gets done not only in classrooms and local programs but in disciplinary contexts — professional organizations, journals and publishers, and accreditation review processes. The work of developing a hospitable architecture extends to organizing these spaces in ways that invite, and perhaps even advocate for, antiracist forms of assessment. Cs, for instance, maintains a position statement ("Writing Assessment," 2022) on writing assessment, detailing core principles for effective writing assessment, documenting the contexts that guide decisions about assessment, and suggesting best practices for assessment, along dimensions from inclusion to learning. What would happen if a similar document were developed, specifically for alternative forms of assessment, to guide teachers and administrators as they sought to establish ungrading in their own programs? Newcomers to ungrading, especially those in faculty roles, are often concerned about accreditation; what would happen if, following Inoue's guidance (2020) on deploying ungrading in established curricula, we as administrators took the lead in mapping out how ungrading may not detract from but strengthen our case for accreditation, supporting educational quality and academic achievement? Until very recently, much of the scholarly work on ungrading has appeared outside flagship journals; what would happen if editorial boards solicited such work, especially from more vulnerable or minoritized scholars? These are only a few examples, yet I hope they serve to illustrate that while I have addressed hostile/hospitable programmatic architecture in the context of writing programs, those architectural or institutional structures may exist in other organized spaces as well. As administrators, who move through, collaborate within, and lead these spaces, we have responsibility to consider how such spaces may be renovated in

ways that invite teachers, especially contingent or otherwise minoritized faculty, to make risky, equitable choices in their assessing and teaching.

Writing researchers

Finally, this work invites writing scholars towards more structurally or institutionally robust conceptualizations of what ungrading is and how the work of ungrading gets done in the classroom. Much research — very good research — on ungrading so far emphasizes individual teachers' choices, within their individual classrooms, and though this work plays an important role in shedding light on how the practice of ungrading works, attending to the multiple, overlapping institutional ecologies that shape ungrading expands and develops our understanding of how ungrading takes root in our programs and comes to bear fruit among our students, in our teaching practices. In this vein, thinking more ecologically, I urge my fellow writing researchers to consider three possible sets of questions. First is the importance of asking students enrolled in ungraded courses about their experience. As I discussed in Chapter 1, I intended to survey students but was unable to complete that work; further research would benefit from centering students' voices, ideally with attention — as I have modelled here — to the influence of institutional architecture or ecologies on students' own experience.

Second and relatedly, the importance of ecological thinking urges research oriented around the relationship between ungrading and stakeholders *besides* students and teachers — administrators, as I have discussed here, but also writing tutors, large-scale assessment companies such as ETS, accrediting bodies such as SACS, upper administration, politicians, and (in K-12 contexts) parents. This work may construct a multi-faceted, ecologically informed sense of how ungrading comes to have the influence it does in the classroom and across the universities and educational institutions we inhabit. Relatedly, I urge fellow researchers to consider how these wider institutional and social contexts shape ungrading. What configurations of programmatic architecture are particularly hospitable to ungrading? Which are particularly hostile? How does geographical or social climate play a role in motivating ungrading? What about political climate?

Finally, I urge fellow researchers to consider where do we want to *go* with our research? How can our work on ungrading be put to good use? Ungrading is not a subject to study with professional detachment; the practice has immediate, real tangible impacts throughout university ecologies, from students and teachers in the classroom to upper administration and to the state level. Considering how we can plan, write, and circulate our research with a view towards its impact in the world(s) that we inhabit is key, an extension of the long tradition of ungrading literature, which from its origins in the 1960s and 70s, staked out a space in the discipline committed to students' and teachers' well-being.

There is more research to be done on ungrading — among other questions, how ungrading is equitable, or inequitable, for neurodivergent and disabled students and faculty, and a keener focus on how writing programs can, even in hostile spaces, make room for people to do this work. I hope subsequent research (my own or others') picks up more fully on the antiracist potential of ungrading, building on this project, and on scholarship such as Inoue's, to explore more fully how specific, culturally sustaining (Paris, 2012) assessment practices gain traction

among writing educators and what institutional supports make actively, audaciously antiracist assessment stick. I like to think that in this project, I have taken an important step in the direction of such questions, laying the groundwork necessary for a rethinking of what ungrading is and how it works to cultivate learning-centered, equitable classrooms that free students from being held to narrow standards for good writing — and how we as teachers, administrators, and scholars, can cultivate environments that help ungrading work well.

A day before writing this last paragraph, I had two fruitful conversations about supporting the work of ungrading — one, with folks holding long-term positions in administration, poised to support the ongoing use of ungrading in first-year composition courses, and another, with a graduate instructor thinking through how ungrading might work in their own contexts, for their own goals. Both of these meetings leave me hopeful. I am not naïve here. I am aware of the great challenges that effective, liberatory forms of higher education face — from political ones, as local, state, and national leaders seek to crush diversity initiatives or defund higher education, to financial ones, and social ones. Against those challenges, however, is persistent interest among the folks who teach courses in the university, and lead the university, in equitable, learning-centered forms of assessment — assessments that work for teachers as well as students, assessments better suited to our own goals and (thoroughly) supported by the programs and departments where we do our work. My hope is that as we offer alternatives to conventional grading, mired in standardized, white American forms of English, we may, as teachers, scholars, and administrators, forge new and more hopeful ways of being and working in first-year writing, in the academe (Royster & Kirsch, 2012). *It doesn't have to be this way,* after all. It can be better — but only if we put in the work. May this project help us to do so.

References

2007 ADE Ad Hoc Committee on Staffing. (2008). *Education in the Balance: A Report on the Academic Workforce in English*. Modern Language Association. https://csal.colostate.edu/docs/cwpa/syllabus/Report-Academic-Workforce.pdf?fbclid=IwAR0hX5xXBnx0vy2sMQm80IHhYJeH9YNI9aYwf8QbwdhRRAEsu2RAWLV40vw

Adsanatham, C. (2012). Integrating Assessment and Instruction: Using Student-Generated Grading Criteria to Evaluate Multimodal Digital Projects. *Computers and Composition*, *29*(2), 152–174. https://doi.org/10.1016/j.compcom.2012.04.002

1804. First Women Admitted. (n.d.). UTK History. https://timeline.utk.edu/history/first-women-admitted

African American Hall of Fame. Office of Multicultural Student Life. https://studentlife.utk.edu/multicultural/history/african-american-hall-of-fame/.

Albracht, L., Harahap, A., Pratt, A., Rodrigo, R., Russell, C., & Schwarz, V. M. (2019). Response to Joyce Olewski Inman and Rebecca A. Powell's "In the Absence of Grades: Dissonance and Desire in Course-Contract Classrooms." *College Composition and Communication*, *71*(1), 145–158.

Background Facts on Contingent Faculty Positions. (n.d.). American Association of University Professors. https://www.aaup.org/issues/contingency/background-facts#:~:text=%22Contingent%E2%80%9D%20faculty%20positions%20include%20both,to%20faculty%20holding%20these%20positions.

Bain, K. (2004). What the Best College Teachers Do. Harvard UP.

Baker-Bell, A. (2021, April 10). *If Black Language Ain't Artful, Then Tell Me, What Is?* [Online presentation.] University of Central Florida, Orlanda, FL.

Baker-Bell, A. (2020). Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy. Routledge.

Barrios, B., Craig, S., & Galin, J. (2023, February 18). *Labor-Based Hope: Three Counternarratives on Complicating Antiracist Grading Practices*. Conference on College Composition and Communication, Chicago, IL.

Belanoff, P. (1991). The Myths of Assessment. *Journal of Basic Writing*, *10*(1), 54–66.

Bennett, K. (2022, October 21). Promoting Usability and Accessibility in Technical & Professional Communication Documents.

Berlin, J. (1987). Rhetoric and Reality: Writing Instruction in American Colleges, 1900—1985. SIU Press.

Birdsall, E. R. (1979). Avoiding Whadjaget with No-Grade, Graded Papers. *College Composition and Communication*, *30*(2), 220. https://doi.org/10.2307/356334

Bisits-Bullen, P. (n.d.). How to: Choose a Sample Size for the Statistically Challenged. *Tools4Dev: Practical Tools for International Development*. https://tools4dev.org/resources/how-to-choose-a-sample-size/

Blum, S. D. (Ed.). (2020). *Ungrading: Why rating students undermines learning (and what to do instead)* (First edition). West Virginia University Press.

Bonilla-Silva, E. (2009). Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America. Rowman & Littlefield.

Boyd, R. (1998). The Origins and Evolution of Grading Student Writing. In F. Zak & C. C. Weaver (Eds.), *Theory and Practice of Grading Writing* (pp. 3–15). SUNY Press.
Brereton, J. (1996). The Origins of Composition Studies in the American College, 1875-1925: A Documentary History. U of Pittsburgh Press.
Broad, B. (2003). *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*. Utah State UP. https://digitalcommons.usu.edu/usupress_pubs/140/
Buchholz, W. J. (1979). Behavioral Evaluation: The Checkmark Grading System. *College Composition and Communication*, *30*(3), 302. https://doi.org/10.2307/356404
Byrd, A. A. (2019). *In Pursuit of an Uncommon Literacy: African American Adults' Experiences in a Computer Code Bootcamp* (Order No. 22589427). Available from ProQuest Dissertations & Theses Global. (2283364490). https://utk.idm.oclc.org/login?url=https://www.proquest.com/dissertations-theses/pursuit-uncommon-literacy-african-american-adults/docview/2283364490/se-2
Carillo, E. C. (2021). The Hidden Inequities in Labor-Based Contract Grading. Utah State UP.
Carlson, J. [@JedMSP.] (18 May 2022) Violence directly linked to misappropriation of research is no longer a hypothetical—it's an actualized harm that should factor into [Tweet.] Twitter, https://twitter.com/JedMSP/status/1526987964466843649.
Carter, R., & Rockson, T. A. (1998). Personalizing the Grading Scale for Beginning Writers. *National Council for Teachers of English*, 125–131.
Carter-Tod, S. (2021, March 29). Rhetoric(s): A Broader Definition. *FEN Blog*. https://compstudiesjournal.com/2021/03/29/rhetorics-a-broader-definition/
Cedillo, C. V. (2017). Diversity, Technology, and Composition: Honoring Students' Multimodal Home Places. *Present Tense*, *6*(2), 1–9.
Chellew, C. (2021, March 15). How Removing Public Seating Creates New Civic Barriers. *Azure*. https://www.azuremagazine.com/article/how-removing-public-seating-creates-new-civic-barriers/
Close, D. (2009). Fair Grades. *Teaching Philosophy*, *32*(4), 361–398.
Colby, R., & Colby, R. S. (2017). Real Faculty But Not: The Full-Time, Nontenure-track Position As Contingent Labor. In S. Kahn, W. Lalicker, & A. Lynch-Biniek (Eds.), *Contingency, Exploitation, and Solidarity: Labor and Action in English Composition* (pp. 91–101). WAC Clearinghouse.
College of Liberal Arts Annual Report, 1936-1937 (?). College of Liberal Arts Annual Reports, AR.0274. University of Tennessee Libraries, Knoxville, Special Collections.
College of Liberal Arts Biennial Report, 1936-1938. College of Liberal Arts Annual Reports, AR.0274. University of Tennessee Libraries, Knoxville, Special Collections.
Connors, R. J. (1981). The Rise and Fall of the Modes of Discourse. *College Composition and Communication*, *32*(4), 444. https://doi.org/10.2307/356607
Connors, R. J. (1997). Composition-Rhetoric: Backgrounds, Theory, and Pedagogy. U of Pittsburgh Press.
Contingent Faculty Positions. (n.d.). American Association of University Professors. https://www.aaup.org/issues/contingency
Cowan, M. (2020). A Legacy of Grading Contracts for Composition. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=150
Craig, S. (2021). Your Contract Grading Ain't It. *WPA Journal*, *44*(3), 145–146.

Creekmore, B. B. (2018). Academic Calendar. In *Volopedia*. University of Tennessee Libraries. https://volopedia.lib.utk.edu/entries/academic-calendar/
Creekmore, B. B. (2018). John Cunyus Hodges. In *Volopedia*. University of Tennessee Libraries. https://volopedia.lib.utk.edu/entries/john-cunyus-hodges/
Crowley, S. (1998). Composition In The University: Historical and Polemical Essays. U of Pittsburgh Press.
Crowley, S. (2006). *Toward a Civil Discourse: Rhetoric and Fundamentalism*. University of Pittsburgh Press.
Danielewicz, J., & Elbow, P. (2009). A Unilateral Grading Contract to Improve Learning and Teaching. *College Composition and Communication*, *61*(2), 244–268.
Davis, A. (2021, April 13). *Color Evasiveness*. https://www.youtube.com/watch?v=6D4aZoqDqOQ&ab_channel=APBSpeakers
Deterding, N. M., & Waters, M. C. (2021). Flexible Coding of In-depth Interviews_ A Twenty-first-century Approach.pdf. *Sociological Methods and Research*, *50*(2). https://doi.org/10.1177/0049124118799377
Dively, R. L. (2015). Standardizing English 101 at Southern Illinois University Carbondale: Reflections on the Promise of Improved GTA Preparation and More Effective Writing Instruction. In M. J. Reiff, A. Bawarshi, M. Ballif, & C. Weisser (Eds.), *Ecologies of Writing Programs: Program Profiles in Context*. Parlor Press.
Duffy, J. (2019). Provocations of Virtue: Rhetoric, Ethics, and the Teaching of Writing. University Press of Colorado.
Edbauer, J. (2005). Unframing Models of Public Distribution: From Rhetorical Situation to Rhetorical Ecologies. *Rhetoric Society Quarterly*, *35*(4), 5–24.
Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, *14*(4), 532–550.
Elbow, P. (1993). Ranking, Evaluating, and Liking: Sorting out Three Forms of Judgment. *College English*, *55*(2), 187. https://doi.org/10.2307/378503
Elbow, P. (1997a). Grading Student Writing: Making it Simpler, Fairer, Clearer. In M. D. Sorcinelli & P. Elbow (Eds.), Writing to Learn: Strategies for Assigning and Responding to Writing Across the Disciplines: New Directions for Teaching and Learning (pp. 127–140). John Wiley & Sons.
Elbow, P. (1997b). Taking time out from grading and evaluating while working in a conventional system. *Assessing Writing*, *4*(1), 5–27. https://doi.org/10.1016/S1075-2935(97)80003-7
Elliot, N. (2005). On a Scale: A Social History of Writing Assessment in America. Peter Lang.
Elliot, N. (2016). A Theory of Ethics for Writing Assessment. *The Journal of Writing Assessment*, *9*(1). http://journalofwritingassessment.org/article.php?article=98
Eyler, J. (2022, March 4). The Call is Coming from Inside the House: How Grades Can Undermine Learning and Jeopardize Our Students' Wellbeing. NEXUS Interdisciplinary Conference, Knoxville, TN.
Fernandes, M. [@magsfern.] (23 Oct 2022) This. I think we are much better off naming the methods for the sake of transparency. Ungrading communicates very little [Tweet.] Twitter, https://twitter.com/magsfern/status/1584271974448390144.

Fishmann, J., & Reiff, M. J. (2015). Taking the High Road: Teaching for Transfer in an FYC Program. In M. J. Reiff, A. Bawarshi, M. Ballif, & C. Weisser (Eds.), *Ecologies of Writing Programs: Program Profiles in Context*. Parlor Press.

Fleming, D., Diab, R., & Shimabukuro, M. (2011). TA Experimentation, 1966-1969. In *From Form to Meaning: Freshman Composition and the Long Sixties, 1957-1974* (pp. 93–132). U Pittsburgh Press.

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 1–20.

Garcia de Müeller, G., & Ruiz, I. (2017). Race, Silence, and Writing Program Administration: A Qualitative Study of US College Writing Programs. *WPA Journal*, *40*(2), 19–39.

Gere, A. R., Curzan, A., Hammond, J. W., Hughes, S., Li, R., Moos, A., Smith, K., Zanen, K. V., & Wheeler, K. L. (2021). Communal Justicing: Writing Assessment, Disciplinary Infrastructure, and the Case for Critical Language Awareness. *College Composition and Communication*, *72*(3), 384–412.

Glenn, C., & Gray, L. (2017). *The Writer's Harbrace Handbook* (6th ed.). Cengage Learning.

Gomes, M., Bellati, B., Hope, M., & LaFerriere, A. (2020). Enabling Meaningful Labor: Narratives of Participation in a Grading Contract. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=152

Gomm, R., Hammersley, M., & Foster, P. (2009). Case Study and Generalization. In R. Gomm, M. Hammersley, & P. Foster, *Case Study Method* (pp. 98–115). SAGE Publications Ltd. https://doi.org/10.4135/9780857024367.d9

Gordon-Reed, A. (2016, September 13). The intense debates surrounding Hamilton don't diminish the musical—They enrich it. *Vox*. https://www.vox.com/the-big-idea/2016/9/13/12894934/hamilton-debates-history-race-politics-literature

Gries, L. (2015). Still Life with Rhetoric: A New Materialist Approach for Visual Rhetorics. Utah State University Press.

Griffith, Allison I, & Smith, Dorothy E. (2014). Introduction. In Griffith, Allison I & Smith, Dorothy E (Eds.), *Under New Public Management: Institutional Ethnographies of Changing Front-Line Work* (pp. 3–22). U of Toronto Press.

Grossoehme, D., & Lipstein, E. (2016). Analyzing longitudinal qualitative data: The application of trajectory and recurrent cross-sectional approaches. *BMC Research Notes*, *9*(1), 136. https://doi.org/10.1186/s13104-016-1954-1

Hawhee, D. (1999). Composition History and the Harbrace College Handbook. *College Composition and Communication*, *50*(3), 504–523. https://doi.org/10.2307/358863

Helton, E. L., & Sommers, J. (2000). Repositioning Revision: A Rhetorical Approach to Grading. *Teaching English in The Two Year College*, *28*(2), 157–164.

Huot, B. (2002). *(Re)Articulating Writing Assessment for Teaching and Learning.pdf*. Utah State UP. https://digitalcommons.usu.edu/usupress_pubs/137/

Inman, J. O., & Powell, R. A. (2018). In the Absence of Grades: Dissonance and Desire in Course-Contract Classrooms. *College Composition and Communication*, *70*(1), 30–56.

Inoue, A. (2007). A reply to Peter Elbow on "Community-Based Assessment Pedagogy." *Assessing Writing*, *12*(3), 228–233. https://doi.org/10.1016/j.asw.2008.01.001

Inoue, A. (2014). Theorizing Failure in US Writing Assessments. *Research in the Teaching of English*, *48*(3), 330–352.

Inoue, A. (2015). *Antiracist Writing Assessment Ecologies: Teaching and Assessing Writing for a Socially Just Future*. WAC Clearinghouse. https://wac.colostate.edu/docs/books/inoue/ecologies.pdf

Inoue, A. (2019a). *Labor-Based Grading Contracts: Building Equity and Inclusion in the Compassionate Writing Classroom*. WAC Clearinghouse. https://wac.colostate.edu/books/perspectives/labor/

Inoue, A. (2019b). How Do We Language So People Stop Killing Each Other, or What Do We Do about White Language Supremacy? *College Composition and Communication*, *71*(2), 352–369.

Inoue, A. (2020, April 11). How Do You Do Labor-Based Grading In Pre-Existing Curricula? *Asao B. Inoue's Infrequent Words*. http://asaobinoue.blogspot.com/2020/04/how-do-you-do-labor-based-grading-in.html

Inoue, A. B. (n.d.). Labor-Based Grading Resources. *Asao B. Inoue's Infrequent Words*. http://asaobinoue.blogspot.com/p/labor-based-grading-contract-resources.html

Johnson, G. P. (2021). Grades as a Technology of Surveillance: Normalization, Control, and Big Data in the Teaching of Writing. In E. Beck & L. Hutchinson Campos (Eds.), *Privacy Matters: Conversations Within and Beyond the Classroom* (pp. 53–72). UP Colorado.

Jordan, J. (1988). Nobody Mean More to Me Than You And the Future Life of Willie Jordan. Harvard Educational Review, 58(3), 363–375. https://doi.org/10.17763/haer.58.3.d171833kp7v732j1

Kennedy, M. M. (1979). Generalizing From Single Case Studies. *Evaluation Quarterly*, *3*(4), 661–678. https://doi.org/10.1177/0193841X7900300409

Klotz, S., & Reardon, K. (2022). Crafting a Writing Response Community Through Contract Grading. *Journal of Response to Writing*, *8*(2), 106–126.

Knapp, J. V. (1976). Contract/Conference Evaluations of Freshman Composition. *College English*, *37*(7), 647. https://doi.org/10.2307/376463

Kryger, K., & Zimmerman, G. X. (2020). Neurodivergence and Intersectionality in Labor-Based Grading Contracts. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=156

Laats, A. (2018). Fundamentalist U: Keeping the Faith in American Higher Education. Oxford UP.

LaFrance, M. (2019). Institutional Ethnography: A Theory of Practice for Writing Studies Researchers. University Press of Colorado.

Lee, R., Ahtone, T., Pearce, M., Goodluck, K., McGhee, G., Leff, C., Lanpher, K., & Salinas, T. (2020). *Morrill Act of 1862 Indigenous Land Parcels Database* [Data set]. High Country News. https://www.landgrabu.org/universities/university-of-tennessee

L'Eplattenier, & Mastrangelo, L. (2004). Historical Studies of Writing Program Administration : Individuals, Communities, and the Formation of a Discipline. Parlor Press.

LeGuin, U.K. (1974). *The Dispossessed*. Harper & Row.

Libertz, Dan. [@DLib24.] (28 April 2022) *Maybe it is the bad taste in my mouth from the Course Hero stuff and those who enable that stuff* [Tweet] https://twitter.com/DLib24/status/1519735864670273537.

Liggett, S., Jordan, K., & Price, S. (2011). Mapping Knowledge-Making in Writing Center Research: A Taxonomy of Methodologies. *The Writing Center Journal*, *31*(2), 50–88.

Litterio, L. M. (2016). Contract Grading in a Technical Writing Classroom: A Case Study. *Journal of Writing Assessment*, *9*(2).
Litterio, L. M. (2018). Contract grading in the technical writing classroom: Blending community-based assessment and self-assessment. *Assessing Writing*, *38*, 1–9. https://doi.org/10.1016/j.asw.2018.06.002
Lockett, A. L., Ruiz, I. D., Sanchez, J. C., & Carter, C. (2021). *Race, Rhetoric, and Research Methods*. WAC Clearinghouse.
MacNealy, M. S. (1999). Strategies for Empirical Research in Writing. Allyn and Bacon.
Mallette, J. C., & Hawks, A. (2020). Building Student Agency Through Contract Grading in Technical Communication. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=158
Matthew, W. (n.d.). *Anti-Racism with Andre Henry*. https://theliturgists.com/anti-racism-with-andre-henry-podcast-page/
McCloud, L. (2023). View of Keeping Receipts: Thoughts on Ungrading from a Black Woman Professor. *Zeal: A Journal for the Liberal Arts*, *1*(2), 101–105.
McDonald, W. U. (1975). Grading Student Writing: A Plea for Change. *College Composition and Communication*, *26*(2), 154–158.
Medina, C., & Walker, K. (2018). Validating the Consequences of a Social Justice Pedagogy. In A. M. Haas & M. F. Eble (Eds.), *Key Theoretical Frameworks* (pp. 46–67). Utah State University Press. https://doi.org/10.7330/9781607327585.c002
Meuwese, M. (2020, August 6). 'Hamilton' ignores the statesman's strategy to fund genocidal warfare against Indigenous Peoples. *The Conversation*. https://theconversation.com/hamilton-ignores-the-statesmans-strategy-to-fund-genocidal-warfare-against-indigenous-peoples-142616
Mishler, E. (1979). Meaning in Context: Is There Any Other Kind? *Harvard Educational Review*, *49*(1), 1–19. https://doi.org/10.17763/haer.49.1.b748n4133677245p
Mobley Finn, K. (2018). What Do GTAs Transfer into Their Pedagogical Practices?: A Qualitative Study about Ph.D. Graduate Teaching Associates and the Transferability of Writing Pedagogy Education. University of Tennessee, Knoxville.
Morgan, Dr. H. A., "The Solution of Present-Day Problems." Correspondence, 1923-1938, Box: 14, Folder: 4. University of Tennessee Presidents' Papers, AR-0001. Betsey B. Creekmore Special Collections and University Archives, University of Tennessee, Knoxville.
Morgan, Dr. H. A., "The State University in Its Public Relations." Correspondence, 1923-1938, Box: 14, Folder: 4. University of Tennessee Presidents' Papers, AR-0001. Betsey B. Creekmore Special Collections and University Archives, University of Tennessee, Knoxville.
Moore, E., Michael, A., & Penick-Parks, M. W. (Eds.). (2017). *The Guide for White Women Who Teach Black Boys*. Corwin.
Moriarty, D., Núñez De Villavicencio, P., Black, L. A., Bustos, M., Cai, H., Mehlenbacher, B., & Mehlenbacher, A. R. (2019). Durable Research, Portable Findings: Rhetorical Methods in Case Study Research. *Technical Communication Quarterly*, *28*(2), 124–136. https://doi.org/10.1080/10572252.2019.1588376

Multiple gradebooks from classes Hodges taught, 1921-1962, IX, Box: 8. John C. Hodges Collection, MS-0401. Betsey B. Creekmore Special Collections and University Archives, University of Tennessee, Knoxville.

Nguyen, C. T. (2022, August 9). Teacher, Bureaucrat, Cop. *The Daily Nous*. https://dailynous.com/2022/08/09/teacher-bureaucrat-cop-guest-post/

Nickoson, L. (2012). Revisiting Teacher Research. In L. Nickoson & M. P. Sheridan (Eds.), *Writing Studies Research in Practice: Methods and Methodologies* (pp. 101–112). SIU Press.

Nilson, L. B. (2014). Specifications Grading: Restoring Rigor, Motivating Students, and Saving Faculty Time. Stylus Publishing.

Nordstrom, G. (2015). Practitioner Inquiry: Articulating a Model for RAD Research in the Writing Center. *The Writing Center Journal*, *35*(1), 87–116.

Original Broadway Cast of Hamilton (2015), *Who Lives, Who Dies, Who Tells Your Story?* https://www.youtube.com/watch?v=_gnypiKNaJE

Pacheco-Vega, R. [@raulpacheco.] (20 Jan. 2023) 2) normative theory and empirics-derived theory both share an element of expectation. Theory establishes what we can expect, and this [Tweet.] Twitter, https://twitter.com/raulpacheco/status/1616577280083431425.

Pacheco-Vega, R. [@raulpacheco] (20 Jan 2023) 3) both normative and empirics-derived theories necessitate IF-THEN-ELSE operators. "IF my theory of social movement formation holds, THEN we should [Tweet.] *Twitter.* https://twitter.com/raulpacheco/status/1616577760360636416.

Paris, D. (2012). Culturally Sustaining Pedagogy: A Needed Change in Stance, Terminology, and Practice. *Educational Researcher*, *41*(3), 93–97. https://doi.org/10.3102/0013189X12441244

Patterson, G. (2015). The Unbearable Weight of Pedagogical Neutrality: LGBTQ issues in the English Studies Classroom. In J. Alexander & J. Rhodes (Eds.), *Sexual Rhetorics: Methods, Identities, Publics* (pp. 134–146). Routledge.

Perryman-Clark, S. M. (2016). Who We Are(n't) Assessing: Racializin Language and Writing Assessment in Writing Program Administration. *College English*, *79*(2), 206–211.

Petty, J. (2016). The London Spikes Controversy: Homelessness, Urban Securitisation and the Question of 'Hostile Architecture.' *International Journal for Crime, Justice and Social Democracy*, *5*(1), 67–81. https://doi.org/10.5204/ijcjsd.v5i1.286

Pimentel, O., Pimentel, C., & Dean, J. (2017). The Myth of the Colorblind Writing Classroom: White Instructors Confront White Privilege in Their Classrooms. In F. Condon & V. A. Young (Eds.), *Performing Antiracist Pedagogy in Rhetoric, Writing, and Communication* (pp. 109–122). https://wac.colostate.edu/docs/books/antiracist/pimentel.pdf

Potts, G. (2010). A Simple Alternative to Grading. *Inquiry: The Journal of the Virginia Community Colleges*, *15*(1), 29–42.

Purdy, D. H. (1975). A Report on an Ungraded Composition Course. *College Composition and Communication*, *26*(2), 204. https://doi.org/10.2307/357119

Quinn, B. (2014, June 13). Anti-homeless spikes are part of a wider phenomenon of "hostile architecture." *The Guardian*. https://www.theguardian.com/artanddesign/2014/jun/13/anti-homeless-spikes-hostile-architecture

Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting White Supremacy in Assessment: Toward a Justice-Oriented, Antiracist Validity Framework. *Educational Assessment*, 1–9. https://doi.org/10.1080/10627197.2022.2042682

Reichert, N. (2003). Practice Makes Perfect: Contracting Quantity and Quality. *Teaching English in the Two-Year College, 31*(1). 60-68.

Royster, J. J., & Kirsch, G. (2012). *Feminist Rhetorical Practices: New Horizons for Rhetoric, Composition, and Literacy Studies*. SIU Press.

Royster, J. J., & Williams, J. C. (1999). History in the Spaces Left. *College Composition and Communication*, *50*(4), 563–584.

Ruddin, L. P. (2006). You Can Generalize Stupid! Social Scientists, Bent Flyvbjerg, and Case Study Methodology. *Qualitative Inquiry*, *12*(4), 797–812. https://doi.org/10.1177/1077800406288622

Ruiz, I. (2016). Reclaiming Composition for Chicano/as and Other Ethnic Minorities: A Critical History and Pedagogy. Palgrave Macmillan.

Sackey, D. J., & DeVoss, D. N. (2012). Ecology, Ecologies, and Institutions: Eco and Composition. In S. Dobrin (Ed.), *Ecology, Writing Studies, and New Media* (pp. 195–212). Routledge.

Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers* (3rd ed.). SAGE publications Ltd.

Sanchez, J. C., & Branson, T. (2016). The Role of Composition Programs in De-Normalizing Whiteness in the University: Programmatic Approaches to Anti-Racist Pedagogies. *WPA Journal*, *39*(2), 47–52.

Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A–F marking scheme. *Journal of Curriculum Studies*, *46*(2), 201–224. https://doi.org/10.1080/00220272.2013.790480

Schwarz, V. M. (2020). (De)Norming Classroom Merit: Grading Contracts as an Assessment Genre. University of Wisconsin, Madison.

Scott, J. C. (1999). Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale UP.

Shellam, T., & Cruickshank, J. (2019). Critical Archives: An introduction. *Journal of Colonialism and Colonial History*, *20*(2). https://doi.org/10.1353/cch.2019.0017

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

Smith, J. K. A. (2016). You Are What You Love: The Spiritual Power of Habit. Brazos Press.

Spidell, C., & Thelin, W. (2006). Not Ready To Let Go: A Study of Resistance to Grading Contracts. *Composition Studies*, *34*(1), 35–68.

Stuckey, M. A., Erdem, E., & Waggoner, Z. (2020). Rebuilding Habits: Assessing the Impact of a Hybrid Learning Contract in Online First-Year Composition Courses. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=159

Test-Optional Pilot Program Continuing. (2021, May 21). University of Tennessee Knoxville News. https://news.utk.edu/2021/05/27/test-optional-applications-continue-through-2025-admission/

Test Score Policy. (n.d.). Undergraduate Admissions. https://admissions.utk.edu/apply/first-year/uts-test-score-policy/#:~:text=The%20University%20of%20Tennessee%20System,to%20Fall%202022%20entering%20students.

Thomas, A. (2017). *The Hate U Give*. Balzer and Bray. 2017.
Tinoco, L., Gage, S., Bliss, A., Baruca, P., Barron, C., & Meyer, C. (2020). Openings, Risks, and Antiracist Futures at a Hispanic-Serving Institution. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=151
UHBristol Clinical Audit Team. (2009). *How to: Set an Audit Sample and Plan Your Data Collection*. University Hospitals Bristol, National Foundation Trust. https://www.uhbristol.nhs.uk/files/nhs-ubht/5%20How%20To%20Sample%20Data%20Collection%20and%20Form%20v3.pdf
University of California, San Diego (2017). writingcensus.ucsd.edu. *National Census of Writing*.
University of Tennessee Manual of Instructions for Freshman English, AR.0764. University of Tennessee Libraries, Knoxville, Special Collections.
University of Tennessee Record. (1926-1927). *University of Tennessee Press*, *29*(3).
Vander Lei, E. (n.d.). Introduction. In E. Vander Lei, T. Amorose, B. Daniell, & A. R. Gere (Eds.), *Renovating Rhetoric in Christian Tradition* (pp. ix–xiv). University of Pittsburgh Press.
Veit, R. C. (1979). De-Grading Composition: Do Papers Need Grades? *College English*, *41*(4), 432–435.
Wardle, E. (2007). Understanding 'Transfer' from FYC: Preliminary Results of a Longitudinal Study. *WPA Journal*, *31*(1–2), 65–85.
Watts, M. (2007). They Have Tied Me to a Stake: Reflections on the Art of Case Study Research. *Qualitative Inquiry*, *13*(2), 204–217. https://doi.org/10.1177/1077800406295628
Wiggins, G. (2014, April 15). What Feedback Is and Isn't. *Granted, and . . . Thoughts on Education by Grant Wiggins*. https://grantwiggins.wordpress.com/2014/04/15/what-feedback-is-and-isnt/
Wilkinson, B. (2005). The Seven Laws of the Learner: How to Teach Almost Anything to Practically Anyone. Multnomah.
Williams, E. (2022, May 2). Do Grades Make Our Lives Worse? *Prindle Post*. https://www.prindleinstitute.org/2022/05/do-grades-make-our-lives-worse/
Wood, S. (2020). Engaging in Resistant Genres as Antiracist Teacher Response. *The Journal of Writing Assessment*, *13*(2). http://journalofwritingassessment.org/article.php?article=157
Wootton, L. (2022). Doing the Right Thing: Professionalism, Deep Accountability, and Emotional Labor in Disciplinary Writing Instruction. George Mason University.
Wootton, L. [@lacey_wootton] (4 April 2022) I used IE as my methodology. I think IE provides useful terms that serve as frames for interpretation. It also, importantly [Tweet.] Twitter, . https://twitter.com/lacey_wootton/status/1511066578636447747?s=20&t=k5fFAj2ew9tfh07B_4y8SQ
Writing Assessment: A Position Statement. (2022). Conference on College Composition and Communication. https://cccc.ncte.org/cccc/resources/positions/writingassessment
Yagelski, R. (1999). Literacy Matters: Writing and Reading the Social Self. Teachers College Press.

Appendices

Appendix A: 2020-2021 Consent Form

Consent for Research Participation
Research Study Title: Contract / Ungrading and GTAs / NTT lecturers at UTK-English
Researchers: Megan Von Bergen, English Department, University of Tennessee, Knoxville; Jeff Ringer, English Department, University of Tennessee, Knoxville

Why am I being asked to be in this research study?
We are asking you to be in this research study because you are a GTA or NTT lecturer in the English department at the University of Tennessee-Knoxville who is using some version of contract / ungrading, whether in part or whole.

While this is an exploratory project, its purpose is to determine what motivates GTAs and contingent faculty to use ungrading (especially in the contexts of both COVID and #BlackLivesMatter); how they learn to use ungrading; and how ungrading works for them as a teacher, especially in light of their teaching philosophy, persona, and identity.

Purpose / Process of the Study
You can expect that if you agree to be in the study, the interview(s) / focus groups will take about two hours at most. You will choose between

- Participating in a focus group, with an optional follow up, one-on-one interview
- Participating in a one-time, one-on-one interview

Focus groups will be held with people at the same level (e.g. all GTAs, or all NTT lecturers), so if there aren't enough other participants at your level, you may be asked if you are willing to switch to doing a one-on-one interview. (You're free to say no if you prefer the focus group.) I expect that participating in a focus group / optional follow-up interview will last no more than 20 minutes longer than participating in a one-time, one-on-one interview. Neither option will last more than about 2 hours.

During the focus group or interview, you will be asked to talk, either on your own or with other GTAs / NTT lecturers from the English department, bout how you adopted contract / ungrading; your motivations for doing so; how it worked for you; and how your experience of contract grading connects with your teaching philosophy and identity.

You will be asked to bring the following materials to the interview:

- Your grading contract / grading matrix / other grading-related document that walks students through your ungrading system and shows how their work translates to a particular grade on the transcript.
- An anonymized document showing the distribution of grades in your class (e.g. # of As, # of Bs, etc).

You may bring a paper or electronic copy to the interview, whichever is more convenient for you. If you bring an electronic copy, simply email it from your myutk address to my myutk

address prior to the start of the focus group / interview. I will retain an electronic copy of the anonymized document with grade distributions, and if you are willing, I would also like to retain an electronic copy of the documents you used to discuss ungrading in your class.

Focus groups and interviews will be held over Zoom, or (if university COVID procedures allow for it) at a park or personal home.

Once your focus group (and optional follow-up interview) or one-on-one interview is done, your participation in the study will be complete.

Opting out of the Study
Being in this study is up to you. You can say no now or leave the study later. Either way, your decision won't affect your grades, your relationship with instructors, or your standing with the composition office, the English department, or the University of Tennessee-Knoxville.

Even if you decide to be in the study now, you can change your mind and stop at any time. Please contact the PI using the information below to opt out of the study.

- Megan Von Bergen, mvonberg@vols.utk.edu

If you opt out before or during the focus group / interviews, your responses will not be transcribed for analysis. Any individual or audiovisual recordings of you will be immediately deleted, though if you participated in a focus group, it will not be possible to delete you from the group recording.

Transcription/analysis is likely to begin no more than six months after the interview is complete. Once transcription begins it is not possible to withdraw your data from the study.

Risks/Benefits from the Study
Risks: It is possible that someone could find out that you were in this study or see your study information if they were to access the audiovisual recording of your interview(s) / focus groups, but we believe this risk is small because of the procedures we use to protect your information. These procedures are described later in this form.

Benefits: You may find that participating in a focus group and/or interview(s) about your use of ungrading gives you the chance to reflect productively on your teaching practices, and make changes for the future, though there is no guarantee this will happen.

Even if you don't benefit from being in the study, your participation may help us to learn more about how GTAs and NTT faculty learn to ungrade and use ungrading in their classes. We hope that in the future, the knowledge gained from this study will benefit WPAs and other writing instructors interested in contract / ungrading.

Privacy & Confidentiality

Your information, including the audiovisual recording of your data and any transcriptions, will be stored on my UTK Google Drive, which is password-protected. Analysis will also be conducted on my password-protected computer. Only me and my faculty advisor will have access to this information.

I will keep your information confidential in presentations, publications, or other published material that results from this study by using a pseudonym, which you will be invite to choose.

The name of the institution will not be redacted, and I may include other information relevant to your teaching (e.g. number of years teaching, status [PhD / 2nd-year MA / NTT lecturer]). While this raises your risk of identification slightly, we believe that 1) using a pseudonym of your choosing, 2) omitting direct identifiers, and 3) describing participants in the plural where appropriate (e.g. if only one male teacher participates in the study, I will write "male teachers" rather than "the one male teacher" when describing connections between gender and ungrading) will prevent anyone who is not on the research team from knowing that you gave us information or what information came from you. Although it is unlikely, there are times when others may need to see the information we collect about you.

We will make every effort to prevent anyone who is not on the research team from knowing that you gave us information or what information came from you. Although it is unlikely, there are times when others may need to see the information we collect about you. These include

- People at the University of Tennessee, Knoxville who oversee research to make sure it is conducted properly.
- Government agencies (such as the Office for Human Research Protections in the U.S. Department of Health and Human Services), and others responsible for watching over the safety, effectiveness and conduct of the research.
- If a law or court requires us to share the information, we would have to follow that law or court final ruling.

Your information will be kept to use for future research and publications, including conference presentations, workshops, posters, articles, chapters, dissertation(s), and books. Your name and other information that can directly identify you will be kept secure and stored separately from your research data collected as part of the study.

We may share your research data with other researchers without asking for your consent again, but it will not contain information that could directly identify you.

What else do I need to know?

This is a small study, with between 3-30 people taking part in this study. Because of the small number of participants, it is possible that someone could identify you based on the information we collected from you, though again, care will be taken to keep your information confidential. (See discussion of privacy and confidentiality procedures above.)

If we learn about any new information that may change your mind about being in the study, we will tell you. If that happens, you may be asked to sign a new consent form.

Questions & Concerns

If you have questions or concerns about this study, or have experienced a research related problem or injury, contact the researchers: Megan Von Bergen (mvonberg@vols.utk.edu) or her advisor, Jeff Ringer (jeff.ringer@utk.edu).

For questions or concerns about your rights or to speak with someone other than the research team about this study, please contact:
Institutional Review Board
The University of Tennessee, Knoxville
1534 White Avenue
Blount Hall, Room 408
Knoxville, TN 37996-1529
Phone: 865-974-7697
Email: utkirb@utk.edu

Statement of Consent

I have read this form, been given the chance to ask questions and have my questions answered. If I have more questions, I have been told who to contact. I understand that I am agreeing to be in this study. I can keep a copy of this consent information for future reference. If I do not want to be in this study, I do not need to do anything else.

I understand that this study will be conducted via the PI's UTK Zoom account and an audiovisual recording will be made. I consent to having the focus group, follow-up interview, and/or one-on-one interview recorded via Zoom, then transcribed and analyzed for research purposes.

[Participants will indicate consent by accessing a Qualtrics survey via CAS authentication (records NetID), then clicking a box to indicate consent, or the lack of consent.]

Appendix B: Interview Protocol 2020-2021

Focus Group

Introductory questions

- Tell me about how contract grading went for you [after our session]? How did you feel about it once the summer / course actually got underway?
- What surprised you / what did you like or dislike?
- What would you say is your big takeaway about ungrading and sustainable, equitable teaching practices?

Logistics. How did it work for you?

One thing I'm interested in is how GTAs [lecturers] learn to ungrade.

- (If participants attended one of my workshops) In the materials I provided you, what resources did you draw on to learn about ungrading and design your approach to ungrading? What additional resources did you use? Tell me about how you used them.
- (If participants did not attend one of my workshops) What resources did you use to learn about ungrading or contract grading? Tell me about how you used them?
- Going forward, what strategies or ways of doing ungrading do you plan to hang onto? Abandon? Revise/adapt/keep?
- In retrospect, what kinds of things were most useful to you from our workshop and the folders I provided? What could you have used more clarification on? What gaps existed? (Alternatively, if participants have not accessed the folders, what kinds of things were most useful to you in the resources you used to adopt contract grading? What kinds of things did you want more clarification on?)

I invited you to bring your **grading contract / grading matrix / other grading-related document(s)** that shows how you framed assessment to students and how their work translated into a transcript grade. Tell me about the choices you made to set it up for you? How would you say it worked for the course?

I also asked you to bring the distribution of grades in your course (how many As, how many Bs, etc). How would you say this distribution is similar to or different from prior courses you've taught? How do you feel about this distribution and why?

Motivations. What made you decide to use ungrading? Did it meet your expectations?
(Goal: I want to know about the reasons or motivations behind people's adoption of contract / ungrading, what they want it to do in the classroom. I want to know specifically how "what I want it to do" is related to COVID; and to anti-racism.)

- What motivated you to make the jump to contract / ungrading? What you were hoping contract / ungrading would do in your writing class? What outcomes were you hoping for?
- Would you say contract / ungrading had the outcomes you were looking for? How so? What outcomes did it have that you did not anticipate?
- How would you say that using ungrading made the course a different experience for you / your students than if you were teaching without using ungrading?

Because COVID. Let's talk (more) about the impact of COVID on your reasons for using ungrading and
its outcomes in your course.

- (If we haven't talked about it yet) Would you say your motivation for switching to ungrading was influenced by COVID? How so? In light of COVID, what outcomes were you hoping for from ungrading?
- (If we have talked about it) You mentioned _____ about COVID; say more about how the pandemic changed your openness to ungrading, or affected the outcomes you were looking for from ungrading?
- How would you say that using contract / ungrading during COVID made the course a better / worse / different experience for you and your students than if you were teaching without ungrading?

Because antiracism. One reason that a lot of people adopt contract / ungrading is because it is an antiracist assessment method.

- Would you say that antiracism, or other structural inequities, factored into your decision to use contract / ungrading? How so?
- Would you say that contract / ungrading had an antiracist impact, or an impact on other structural inequities, in your class? How so?

Personal experience(s): You are all PhD GTAs / MA GTAs / lecturers.

- How would you say that being a GTA or NTT lecturer affected your decision to use contract grading and the way that it turned out for you?
- Would you say that being a GTA / NTT lecturer made your experience with ungrading different / harder / easier, compared with somebody at a different level? How so?
- How would you say that ungrading changed your stress levels, or how you feel about teaching and grading generally?

Wrap Up

Is there anything else y'all would like me to know about your use of ungrading / contract grading?

Focus Group (Individualized Follow Up)

Here's what I really want to do w/ this: I want it to be a chance to ask more specific questions about how people learned to ungrade / how it changed their teaching philosophy / what worked and didn't work for them as a TA that I couldn't ask elsewhere. I may skip questions or sections,

depending on whether participants' original answers in the focus group are sufficient (e.g. if a participants' responses during the focus group answered the covid questions fully, I may skip it altogether.)

Follow up(s)

- You said _____. Can you tell me more about that?
- (Repeat as needed for other insightful or interesting things that participants said. This portion of the follow-up will focus on drawing out more information about things participants have already said; I will not ask for new information.)

Covid

I know that for a lot of people, COVID provided the exigence to get into contract / ungrading, so I want to come back to that and talk a little more about your personal experience using ungrading during the pandemic.

- You said _____ about using ungrading during COVID. Can you tell me more about that? In terms of teaching during COVID, what has been most difficult for you? Has ungrading helped with that at all? How so?
- How would you describe your goals as a teacher during COVID? Has ungrading helped you meet these goals? How so?
- Is there anything else you would like to share with me about using ungrading during COVID?

Antiracism

#BlackLivesMatter also provided an exigence for people to start ungrading, so I want to talk in a little more detail about your personal experience with that.

- When we talked about anti-racism / structural inequities, and ungrading, you said
- _____. Can you say more about that?
- How would you describe your goals as a teacher regarding antiracism, or structural inequities?
- Has ungrading helped you meet these goals?
- If you did not have concrete goals, has ungrading helped you rethink what your goals are? How so?
- Is there anything else you would like to share with me about using ungrading and anti-racism / structural inequities?

Experience of teaching/grading

Last, I want to talk briefly about how ungrading fits in with your experience and beliefs about teaching, along with who you are as a person.

- (If participants have 3+ semesters of independent teaching experience) By now, you have a fairly good sense of who you are as a teacher; how did ungrading go along with that, or change that? How does ungrading fit in with what you believe about teaching or change it?

- (If participants have 2 or fewer semesters of independent teaching experience) How does ungrading go along with the kind of teachers you want to become? How would you say that ungrading fits in with what you believe about teaching?
- Prior to using ungrading, how did you feel about grading or assessing students' work? How did using ungrading change that feeling, if at all?
- You are a (man/woman/nonbinary/genderqueer). How do you think your gender identity affects your teaching persona in ways that make it easier / harder / different for you to adopt ungrading, compared with somebody of a different gender?

As you know, my goals with this study are 1) to find out what motivates GTAs / lecturers to adopt contract grading and 2) how they learn to do it. Is there anything that you want to say now, that you didn't get a chance to say in our first session or that you want to elaborate on from our workshop?

Individualized Interviews (No focus group option)

Introductory questions

- Tell me about how contract grading went for you after my workshop / after you actually put it into place? How did you feel about it once the summer / course actually got underway?
- What was your big takeaway about the connections between ungrading and equitable, sustainable ways of teaching?

Logistical questions

One thing I'm interested in is how GTAs [lecturers] learn to ungrade.

- (If participants attended one of my workshops) In the materials I provided you, what resources did you draw on to learn about ungrading and design your approach to ungrading? What additional resources did you use? Tell me about how you used them.
- (If participants did not attend one of my workshops) What resources did you use to learn about ungrading or contract grading? Tell me about how you used them?
- Going forward, what strategies or ways of doing ungrading do you plan to hang onto? Abandon? Revise/adapt/keep?
- In retrospect, what kinds of things were most useful to you from our workshop and the folders I provided? What could you have used more clarification on? What gaps existed? (Alternatively, if participants have not accessed the folders, what kinds of things were most useful to you in the resources you used to adopt contract grading? What kinds of things did you want more clarification on?)

I invited you to bring your grading contract / grading matrix / other grading-related document that shows how you framed assessment to students and how their work translated into a transcript grade. Tell me about the choices you made to set it up? How would you say it worked for the course?

I also asked you to bring the distribution of grades in your course (how many As, how many Bs, etc). How would you say this distribution is similar to or different from prior courses you've taught? How do you feel about this distribution?

Motivations. What made you decide to use ungrading? Did it meet your expectations?

(Goal: I want to know about the reasons or motivations behind people's adoption of contract / ungrading, what they want it to do in the classroom. I want to know specifically how "what I want it to do" is related to COVID; and to anti-racism.)

- What motivated you to make the jump to contract / ungrading? What you were hoping contract / ungrading would do in your writing class? What outcomes were you hoping for?
- Would you say contract / ungrading had the outcomes you were looking for? How so? What outcomes did it have that you did not anticipate?
- How would you say that using ungrading made the course a different experience for you / your students than if you were teaching without using ungrading?

Because COVID. Let's talk (more) about the impact of COVID on your reasons for using ungrading and its outcomes in your course.

- (If we haven't talked about it yet) Would you say that your pedagogical goals during COVID motivated your switch to ungrading? How so? Did ungrading help you meet these goals or not? How so?
- (If we have talked about it) You mentioned _____ about COVID; say more about how your openness to ungrading, or affected the outcomes you were looking for from ungrading / the pandemic helped you meet your goals as a teacher during COVID?
- How would you say that using contract / ungrading during COVID made the course a better / worse / different experience for you and your students?

Because antiracism. One reason that a lot of people adopt contract / ungrading is because it is an antiracist assessment method.

- (If we haven't talked about it) Would you say that your teaching goals regarding antiracism / structural inequities played a role in your switch to ungrading? How so? (If you did not have concrete goals, has using ungrading helped you think about what your goals are relative to antiracism / structural inequities? How so?) How has ungrading helped you meet these goals? How has ungrading had an antiracist impact / impact on other structural inequities in your class?
- (If we've talked about it) You mentioned _____ about antiracism. Can you say more about how ungrading had an antiracist impact / impact on other structural inequities in your class?

Personal experience:

- How would you say that being a TA or NTT lecturer affected your decision to use contract grading and the way that it turned out for you? Would you say that being a GTA / NTT lecturer made your experience with ungrading different / harder / easier, compared with somebody at a different level? How so?

- (If participants have 3+ semesters of independent teaching experience) By now, you have a fairly good sense of who you are as a teacher; how did ungrading go along with that, or change that?
- (If participants have 2 or fewer semesters of independent teaching experience) How does ungrading go along with the kind of teachers you want to become?
- You are a (man/woman/nonbinary/genderqueer). How do you think your gender identity affects your teaching persona in ways that make it easier / harder / different for you to adopt ungrading, compared with somebody of a different gender?
- Prior to using ungrading, how did you feel about grading or assessing students' work? How did using ungrading change that feeling, if at all?

Wrap Up

As you know, my goals with this study are 1) to find out what motivates GTAs / lecturers to adopt contract grading and 2) how they learn to do it. Is there anything else y'all would like me to know about your use of contract grading / ungrading?

Appendix C: 2021-2022 consent form

Consent for Research Participation — Focus Group / Follow-Up Interview
Title: Ungrading in the UTK First-Year Composition Program
Researchers: Megan Von Bergen, Dr. Jeff Ringer, English Dept, U. of Tennessee, Knoxville

We are asking you to be in this research study because you are a GTA or NTT lecturer in the FYC ungrading pilot during Summer 2021, Fall 2021, and/or Spring 2022.

The goal of this study is to determine how ungrading, with the active support of ungrading by a writing program, changes what students learn in first-year writing courses and how teachers and students experience the course.

Study Processes
You will choose between

- Participating in a focus group, with an optional follow up, one-on-one interview
- Participating in a one-time, one-on-one interview

If, when you opt in, there are not enough other participants for a focus group, you may be asked if you are willing to switch to doing a one-on-one interview. (You're free to say no if you prefer the focus group.) Regardless of which option you choose, the interview and/or focus group will not last more than 2-3 hours.

During the focus group or interview, you will be asked to talk about your motivations for adopting ungrading, how you implemented it in your ENGL 101 / 102 / 131 / 132 class and its effects on your teaching, and how your experience of contract grading connects with your teaching philosophy and identity.

You will provide your grading contract / grading matrix / other grading-related documents and an anonymized document showing the distribution of grades in your class (e.g. # of As, # of Bs, etc). I will keep copies of these documents for analysis.

Focus groups and interviews will be held over Zoom, or (if university COVID procedures allow for it) at a park or personal home.

Once your focus group (and optional follow-up interview) or one-on-one interview is done, your participation in the study will be complete.

Opting out of the Study

Being in this study is up to you. You can say no now or leave the study later. Either way, your decision will have zero effect on your ability to participate in the pilot study, your standing with the First-Year Composition office, the English department, or the University of Tennessee-Knoxville.

Even if you decide to be in the study now, you can change your mind and stop at any time. Please contact Megan Von Bergen (mvonberg@vols.utk.edu) to opt out of the study.

If you opt out before or during the focus group / interviews, your responses will not be transcribed for analysis. Any individual or audiovisual recordings of you will be immediately deleted, though if you participated in a focus group, it will not be possible to delete you from the group recording.

Transcription/analysis is likely to begin no more than two months after the interview is complete. Once transcription begins it is not possible to withdraw your data from the study.

Risks/Benefits from the Study

Risks: As with all research, there is a risk of a breach of confidentiality, but I believe this risk is small because of the steps we take to protect your information. These procedures are described later in this form.

Benefits: You may find that participating in a focus group and/or interview(s) about your use of ungrading gives you the chance to reflect productively on your teaching practices, and make changes for the future, though there is no guarantee this will happen.

Your participation may help us to learn more about how GTAs and NTT faculty use ungrading and how writing programs / WPAs can support instructors in using ungrading and alternative assessment methods.

Privacy & Confidentiality

To ensure privacy, I will 1) store all research material on my UTK G'Drive and analyze it on my personal, password-protected computer, 2) I will be the only person with access to the video data (my faculty advisors will have access to anonymized transcripts), and 3) your data will be assigned a pseudonym, which I will use in presentations, publications, or other published material resulting from this study.

The name of the institution will not be redacted, and I may include other information relevant to your teaching (e.g. number of years teaching, status [PhD / 2nd-year MA / NTT lecturer]). While this raises your risk of identification slightly, we believe that 1) using a pseudonym of your choosing, 2) omitting direct identifiers, and 3) describing participants in the plural where appropriate (e.g. if only one male teacher participates in the study, I will write "male teachers" rather than "the one male teacher" when describing connections between gender and ungrading) will prevent anyone who is not on the research team from knowing that you gave us information or what information came from you. Although it is unlikely, there are times when others may need to see the information we collect about you.

Although it is unlikely, there are times when others may need to see the information we collect about you. These include

- People at the University of Tennessee, Knoxville who oversee research to make sure it is conducted properly.
- Government agencies (such as the Office for Human Research Protections in the U.S. Department of Health and Human Services), and others responsible for watching over the safety, effectiveness and conduct of the research.
- If a law or court requires us to share the information, we would have to follow that law or court final ruling.

Your information will be kept to use for future research and publications, including conference presentations, workshops, posters, articles, chapters, dissertation(s), and books. Your name and other information that can directly identify you will be kept secure and stored separately from your research data collected as part of the study.

We may share your research data with other researchers without asking for your consent again, but it will not contain information that could directly identify you.

Will I be paid for being in this study?

Yes! You will receive a $25 gift card for participating in a focus group, and a $20 gift card for participating in an interview. Gift cards will be for one of three local places: Wild Love, Union Books, or Yassins — your choice!

What else do I need to know?
This is a small study, with no more than 50 instructors taking part in it. Because of the small number of participants, it is possible that someone could identify you based on the information we collected from you, though again, care will be taken to keep your information confidential. (See discussion of privacy and confidentiality procedures above.)

If we learn about any new information that may change your mind about being in the study, we will tell you. If that happens, you may be asked to sign a new consent form.

Questions & Concerns

If you have questions or concerns about this study, or have experienced a research related problem or injury, contact the researchers: Megan Von Bergen (mvonberg@vols.utk.edu) or her advisor, Jeff Ringer (jeff.ringer@utk.edu).

For questions or concerns about your rights or to speak with someone other than the research team about this study, please contact:
Institutional Review Board
The University of Tennessee, Knoxville
1534 White Avenue
Blount Hall, Room 408
Knoxville, TN 37996-1529
Phone: 865-974-7697

Email: utkirb@utk.edu

Statement of Consent
I have read this form and I agree to participate in this study.
I consent to having an audiovisual recording of the focus group and/or any interview made via Zoom.

[Participants will indicate consent by accessing a Qualtrics survey via CAS authentication (records NetID), then clicking a box to indicate consent, or the lack of consent.]

Appendix D: 2021-2022 Recruitment

Hi!

I'm reaching out to you because you're participating in the First-Year Composition program's pilot on ungrading. I hope it's going great for you, and I have a follow-up question — would you be interested in participating in research on ungrading?

As you may know, my dissertation focuses on ungrading in first-year composition courses. I want to know 1) how writing programs can actively support instructors who choose to ungrade and 2) how ungrading changes the experience of students and teachers in first-year composition programs. I'm hoping you'd be willing to share with me about your experiences using ungrading: you'd join a focus group, then (if you're willing) an optional, follow-up interview; each one lasts about an hour. Your viewpoint could help me think more deeply about how writing programs can design and maintain equitable assessment programs that work for both students and teachers. You'd also send a link to a survey for students to your first-year writing class (I'll provide text for the email).

If you're up for it, click here [link to Qualtrics survey] to consent. The link will ask for your netID and password (so we know it's you) and will provide more detailed info about the study.

Thank you, and let me know if you have any questions?

Megan Von Bergen

Follow up

Hi!

Just wanted to follow up on my previous email, about participating in research on ungrading in first-year composition courses. If you are interested and have the time, I'd love to have your input — it's important to me to hear from writing instructors as I pull my research together.

You can use the link [link here to Qualtrics survey] to consent. The link will ask for your netID and password (so we know it's you) and will provide more detailed info about the study.

If you have any questions, feel free to reach out!

All the best, Megan

Appendix E: 2021-2022 Interview Protocols

SEMISTRUCTURED INTERVIEW PROTOCOL — INSTRUCTORS

Procedures

When people are recruited into the study, they will get a choice between a focus group w/ follow up, or an individual interview, though I expect both options to take about the same amount of time. Note that depending on participants' answers, I may skip questions (e.g. if participants fully address COVID's impact on their teaching early on, there may be no need to ask explicitly about it later in the interview, or in the individualized follow up).

Focus groups will include between 2-6 people in them. If there are insufficient people to form a focus group, participants will do an individualized interview. If a focus group is formed, however, and participants opt into it, they will have the choice to participate in just the focus group or in the focus group and individualized interview. These options will be negotiated after recruitment.

Consent forms will be attached to each recruitment letter.
Focus Group
Focus Group (Individualized Follow Up)
Individualized Interviews (No focus group option)

Focus Group

Introductory questions

- Tell me about how contract grading is going for you / went for you?
- What surprised you / what did you like or dislike?
- What would you say is your big takeaway about ungrading and sustainable, equitable teaching practices?

Logistics. How did it work for you?

Before the interview, you submitted your grading contract / matrix / grading related documents, showing how you framed assessment to students and how their work translated into a transcript grade.

Using your course documents as a guide, tell me about the strategies or ways of doing ungrading that you used this year. What do you plan to hang onto? Abandon/revise/adapt? If you've used ungrading before, in what way(s) has your ungrading practice changed?

Also, tell me about how the pilot helped get you there:

- What was useful for you in the pilot / what did you like?
- What was less useful for you in the pilot / what did you not like?
- What would you have liked more training on?

- What resources (if any) did you use outside of our training to set up ungrading? How did it work for you?
- How did having the pilot program in place affect your decision to ungrade / your use of
- ungrading / your ongoing ungrading practice?

I also asked you to share **the distribution of grades in your course** (how many As, how many Bs, etc). How would you say this distribution is similar to or different from prior courses you've taught, whether conventionally graded or ungraded? How do you feel about this distribution and why?

Motivations. What made you decide to use ungrading? Did it meet your expectations?

- (If participants are new to ungrading) Tell me about why you made the jump to ungrading. What motivations factored into your decision to use it? What outcomes did you anticipate / what outcomes did it actually have?
- (If participants have used ungrading in the past) You've used ungrading before; what motivations factored into your decision to keep using ungrading? How so? What outcomes did you anticipate / what outcomes did it actually have?
- How do you see the function of ungrading in your classroom, for you, your students, and your teaching? What reservations did you have about ungrading, and what has happened to them? Has your idea about what ungrading is and what it does for you, your students, and/or your teaching changed, the longer you've used ungrading? How so?
- How would you say that using ungrading made the course a different experience for you / your students than if you were teaching without using ungrading? How do you think the course remained the same?

Because antiracism. One reason that a lot of people adopt contract / ungrading is because it is an anti-racist assessment method.

- Would you describe your use of ungrading as antiracist? Why or why not?
- In what way(s) do you think your use of ungrading has antiracist outcomes? How do you think using ungrading made your course more / less / different for raciolinguistically diverse student than if you were not using ungrading? How have your ideas about ungrading as antiracist / your use of ungrading for antiracist purposes changed, the longer you've been using ungrading?
- What are some gaps in the antiracist work of your ungrading practice?
- How could the writing program / pilot support you in making your ungrading practice more antiracist?

Because COVID. Let's talk (more) about the impact of COVID on your reasons for using ungrading and its
outcomes in your course.

- (If we haven't talked about it yet) How would you say that COVID factored into your decision to use ungrading / your decision to keep using ungrading? In light of COVID, what were you hoping ungrading would do in your course?

- (If we have talked about it) You mentioned _____ about COVID; say more about how COVID factored into your decision to use ungrading / your decision to keep using ungrading?
- How would you say that ungrading made the course better / worse / different for you and your students than if you were teaching without ungrading during COVID?
- How could the writing program support you in using ungrading more during the pandemic / social crises?

Personal experience(s)
- How would you say that being a GTA or NTT lecturer affected your decision to use contract grading and the way that it turned out for you?
- Would you say that being a GTA / NTT lecturer made your experience with ungrading different / harder / easier, compared with somebody at a different level? How so?
- How would you say that ungrading changed your stress levels, or how you feel about teaching and grading generally?

Wrap Up
Is there anything else y'all would like me to know about your use of ungrading / contract grading?

Focus Group (Individualized Follow Up)

Note: I may skip questions or sections, depending on whether participants' original answers in the focus group are sufficient (e.g. if a participants' responses during the focus group answered the covid questions fully, I may skip it altogether.)

Follow up(s)
- You said _____. Can you tell me more about that?
- (Repeat as needed for other insightful or interesting things that participants said. This portion of the follow-up will focus on drawing out more information about things participants have already said.)

(If participants have used ungrading for more than one semester)
- Would you say that your ideas about ungrading / ungrading practice has changed, compared to where you started? How so?
- You said _____ about the way your ungrading philosophy / practice changed in our focus group. Can you tell me more about that?

Antiracism
The urgent need for antiracist assessment also provided an exigence for people to start ungrading, so I want to talk in a little more detail about your personal experience with that.
- When we talked about antiracism / structural inequities, and ungrading, you said _____. Can you say more about that?
- How would you describe your goals as a teacher regarding antiracism, or structural inequities?

- Has ungrading helped you meet these goals? If you did not have concrete goals, has ungrading helped you rethink what your goals are? How so?
- You've been ungrading for _____ semesters now. How have your ideas about ungrading as antiracist / practice of ungrading as antiracist changed during this time?
- What "gaps" in your ungrading practice do you want to address to make it more antiracist? What do you want to keep learning about ungrading and antiracism?
- Is there anything else you would like to share with me about using ungrading and anti-racism / structural inequities?

Covid

I know that for a lot of people, COVID provided the exigence to get into contract / ungrading, so I want to come back to that and talk a little more about your personal experience using ungrading during the pandemic.

- You said _____ about using ungrading during COVID. Can you tell me more about that?
- In terms of teaching during COVID, what has been most difficult for you? Has ungrading helped with that at all? How so?
- You've been ungrading for _____ semesters now. How have your ideas about ungrading as an assessment response to the pandemic/social crises changed during this time?
- How would you describe your goals as a teacher during COVID? Has ungrading helped you meet these goals? How so?
- Is there anything else you would like to share with me about using ungrading during COVID?

Experience of teaching/grading

Last, I want to talk briefly about how ungrading fits in with your experience and beliefs about teaching, along with who you are as a person.

- (If participants have 3+ semesters of independent teaching experience) By now, you have a fairly good sense of who you are as a teacher. In what way(s) has ungrading changed or affirmed (or a bit of both) your teaching philosophy / identity as a teacher? If you've used ungrading previously, how has your teaching philosophy / identity continued to change, the longer you've been using it?
- (If participants have 2 or fewer semesters of independent teaching experience) How does
- ungrading go along with the kind of teacher you want to become? How would you say that ungrading fits in with what you believe about teaching? If you've used ungrading previously, how has your teaching philosophy / identity continued to change, the longer you've been using it?
- Prior to using ungrading, how did you feel about grading or assessing students' work? How did using ungrading change that feeling, if at all?
- You are a (man/woman/nonbinary/genderqueer). How do you think your gender identity affects your teaching persona in ways that make it easier / harder / different for you to adopt ungrading, compared with somebody of a different gender?

- You are (white/Black/Asian/Latinx/Indigenous/Pacific Islander/biracial/multiracial). How do you think your racial identity affects your use of ungrading in ways that make it easier / harder / different for you to adopt the practice and convince your students to get on board, compared with somebody of a different race?

As you know, my goals with this study are to find out 1) how using ungrading changes how people teach and how they value writing / writing education and 2) how receiving active support from the writing program, in the context of the pilot program, affects how people implement ungrading, or how they learn to do it. Is there anything that you want to say now, that you didn't get a chance to say in our first session or that you want to elaborate on from our workshop?

Individualized Interviews (No focus group option)

Introductory questions

- Tell me about how contract grading is going for you / went for you?
- What surprised you / what did you like or dislike?
- What would you say is your big takeaway about ungrading and sustainable, equitable teaching practices?

Logistics. How did it work for you?

Prior to the interview, you submitted your grading contract / matrix / grading related documents showing how you framed assessment to students and how their work translated into a transcript grade. Using your course documents as a guide, tell me about the strategies or ways of doing ungrading that you used this year. What do you plan to hang onto? Abandon/revise/adapt? If you've used ungrading before, in what way(s) has your ungrading practice changed?

Also, tell me about how the pilot helped get you there:

- What was useful for you in the pilot / what did you like?
- What was less useful for you in the pilot / what did you not like?
- What would you have liked more training on?
- What resources (if any) did you use outside of our training to set up ungrading? How did it work for you?
- How did having the pilot program in place affect your decision to ungrade / your use of
- ungrading / your ongoing ungrading practice?

I also asked you to bring **the distribution of grades in your course** (how many As, how many Bs, etc). How would you say this distribution is similar to or different from prior courses you've taught, whether conventionally graded or ungraded? How do you feel about this distribution and why?

Motivations. What made you decide to use ungrading? Did it meet your expectations?

- (If participants are new to ungrading) Tell me about why you made the jump to ungrading. What motivations factored into your decision to use it? What outcomes did you anticipate / what outcomes did it actually have?

- (If participants have used ungrading in the past) You've used ungrading before; what motivations factored into your decision to keep using ungrading? How so? What outcomes did you anticipate / what outcomes did it actually have?
- How do you see the function of ungrading in your classroom, for you, your students, and your teaching? What reservations did you have about ungrading, and what has happened to them?
- Has your idea about what ungrading is and what it does for you, your students, and/or your teaching changed, the longer you've used ungrading? How so?
- How would you say that using ungrading made the course a different experience for you / your students than if you were teaching without using ungrading? How do you think the course remained the same?

Because antiracism. One reason that a lot of people adopt contract / ungrading is because it is an anti-racist assessment method.

- Would you describe your use of ungrading as antiracist? Why or why not?
- In what way(s) do you think your use of ungrading has antiracist outcomes? How do you think using ungrading made your course more / less / different for raciolinguistically diverse students than if you were not using ungrading? How have your ideas about ungrading as antiracist / your use of ungrading for antiracist purposes changed, the longer you've been using ungrading?
- You've been ungrading for _____ semesters now. How have your ideas about ungrading as antiracist / practice of ungrading as antiracist changed during this time?
- What are some gaps in the antiracist work of your ungrading practice?
- How could the writing program / pilot support you in making your ungrading practice more antiracist?

Because COVID. Let's talk (more) about the impact of COVID on your reasons for using ungrading and its
outcomes in your course.

- (If we haven't talked about it yet) How would you say that COVID factored into your decision to use ungrading / your decision to keep using ungrading? In light of COVID, what were you hoping ungrading would do in your course / for you / for your students?
- (If we have talked about it) You mentioned _____ about COVID; say more about how COVID factored into your decision to use ungrading / your decision to keep using ungrading?
- How would you say that ungrading made the course better / worse / different for you and your students than if you were teaching without ungrading?
- You've been ungrading for _____ semesters now. How have your ideas about ungrading as an assessment response to the pandemic / social crises changed during this time?
- How could the writing program support you in using ungrading more during the pandemic / social crises?

Personal Experience of Ungrading

Last, I want to talk briefly about how ungrading fits in with your experience and beliefs about teaching, along with who you are as a person.

Personal experience(s)

- How would you say that being a GTA or NTT lecturer affected your decision to use contract grading and the way that it turned out for you?
- Would you say that being a GTA / NTT lecturer made your experience with ungrading different / harder / easier, compared with somebody at a different level? How so?
- How would you say that ungrading changed your stress levels, or how you feel about teaching and grading generally?

- (If participants have 3+ semesters of independent teaching experience) By now, you have a fairly good sense of who you are as a teacher. In what way(s) has ungrading changed or affirmed (or a bit of both) your teaching philosophy / identity as a teacher? If you've used ungrading previously, how has your teaching philosophy / identity continued to change, the longer you've been using it?

- (If participants have 2 or fewer semesters of independent teaching experience) How does ungrading go along with the kind of teacher you want to become? How would you say that ungrading fits in with what you believe about teaching? If you've used ungrading previously, how has your teaching philosophy / identity continued to change, the longer you've been using it?
- Prior to using ungrading, how did you feel about grading or assessing students' work? How did using ungrading change that feeling, if at all?
- You are a (man/woman/nonbinary/genderqueer). How do you think your gender identity affects your teaching persona in ways that make it easier / harder / different for you to adopt ungrading, compared with somebody of a different gender?
- You are (white/Black/Asian/Latinx/Indigenous/Pacific Islander/biracial/multiracial). How do you think your racial identity affects your use of ungrading in ways that make it easier / harder / different for you to adopt the practice and convince your students to get on board, compared with somebody of a different race?

As you know, my goals with this study are to find out 1) how using ungrading changes how people teach and how they value writing / writing education and 2) how receiving active support from the writing program, in the context of the pilot program, affects how people implement ungrading, or how they learn to do it. Is there anything that you want to say about this?

Appendix F: Collection and Audit of Archival Data

This appendix includes my methods for recording and auditing the grade data included in Chapter 2 of this dissertation. Because that chapter uses significant amounts of quantitative data, I conducted an audit to ensure the data was as accurate as possible. I would also like to note that to ensure that all data remains as secure as possible, much of it is presented in the aggregate (e.g. it is not individual records). The few individual records I do address, I address without their name, using a pseudonym where needed to protect individual identity).

In Chapter 2, I include grade distributions for the sections of ENGL 111, 112, and 113 that Hodges taught between AY 1926-27, when the new curriculum was introduced, and AY 1936-37. I created a Google Sheets document for each academic year that Hodges taught, with a separate page for each course. I entered students' names in one column, and in the adjoining column, I entered the quarterly grade that students received, including any pluses and minuses. I added up the total number of each grade assigned per section, the total number of each grade assigned over the academic year, and the total number of students enrolled. While pluses and minuses were used at the time and appear in Hodges's gradebook, I recorded only the letter grade, to provide a clearer, big-picture view of the grade distribution in the course overall.
If a student's name was entered as part of the roster and then crossed out, I entered that name for the sake of comprehensiveness and recorded the grade as "Other." In some cases, there was a note that the student transferred to another section, but these students were also included as "Other" grades, largely because it was impossible to distinguish consistently between students who transferred between sections and students who received an incomplete; the "Other" grade also includes students whose final quarterly grade is marked with an X, blacked out, or otherwise recorded in such a way that the student does not appear to have received a conventional letter grade. If students' names were entered at the bottom of the gradebook, apart from the regular roster, with no information about assignments completed, I omitted those students from my tally. Students who repeat courses — for instance, a student who passes ENGL 111 then fails ENGL 112 and re-takes it the following term — are recorded as two separate students, so that the number of students recorded is not identical with the number of different human beings who sat in Hodges's courses.

I collected much of this data in March 2020, the day before the pandemic closed access to our Special Collections for months; the rest of the data was collected before September 2021, during a still-chaotic period in higher education and research. Also, centuries-old quantitative data is often difficult to read and record accurately. In light of this, I audited my data before finishing analysis. The audit was conducted in November 2021, after I finished a draft of this chapter but before the draft was shared with my committee. The audit followed the process described below: I have twelve Google Sheets documents recording the grades given in any particular year. One sheet is a duplicate, so there are eleven distinct years, with one year (1935-1936) including fewer courses taught by John C. Hodges. Using the total number of students enrolled in each course recorded in each Sheet document, I used my phone calculator to determine that there were approximately 1149 students enrolled over the entire period analyzed; this is an average of 104 students, or 10% of the enrollment in each year (this varies slightly, as enrollment increases in the 1930s). I decided that I would choose sections for the audit at random, eliminating 35-36 from the pool of possible courses since it deviated from the norm.

Using the Google random number generator, I selected the following years:

1. 1929-1930 (when selected, this year turned out to only have one section so I selected another year at random and do the second section for that year)
2. 1930-1931
3. 1926-1927
4. 1934-1935 (second sections only)

I set my goal error rate below 5%. If the error rate is above 5%, additional work would be done to ensure that data and grade distributions was accurate.

I conducted the audit on 18 November 2021. To conduct the audit, I pulled the box with the gradebooks in them, then the gradebooks identified through the random selection above. I created new spreadsheets for each year audited. Each spreadsheet included the following information: 1) student name, 2) grade originally entered, 3) grade entered during audit, and 4) error/no error. I started by entering the student names and the grade entered during audit, working from the original gradebooks; I then entered the grades from the original spreadsheets, and I compared the two. If I located an error, which I did several times, I checked the original spreadsheet against the gradebook. In a few cases, checking the original spreadsheet against the gradebook surfaced an error in the audit data entry, so I corrected the error and then compared the grades from the original data entry and the audit. If the two were in alignment, then I recorded "No Error." If, however, checking the original spreadsheet against the gradebook indicated an error in the *original* data entry which formed the basis for the tables in this document, I recorded an "Error."

I examined 360 records overall and identified 7 errors. All 7 errors appeared in the 1926-1927 dataset. Since this dataset was the first one I created, the higher error rate is unsurprising and speaks to the fact that I had not yet developed a consistent methodology at the time. Of the errors located, only one error was substantial: a B+ that I inadvertently entered as a C+. Three errors were transposed: a D, B/B-, and an I were entered in both the original spreadsheet and the audit but in a different order; since my analysis does not account for pluses/minuses, no errors were reflected in the tables included in this data. Finally, 3 more errors were insubstantial, meaning that only the plus or minus changed, likely due to Google Sheets' autofill. Again, because pluses and minuses were not reflected in the dataset, these errors do not affect the outcomes. Ultimately, the audit had a 2% error rate overall, with a 6% error rate (.058) within 1926-1927; if only the substantial error is included, however, the error rate drops to .008. The audit was successful, and it was determined that further work to confirm the accuracy of data is not needed at this time.

Appendix G: Grade Distributions, 1928-1929 through 1936-1937

For the sake of clarity, I included a histogram representing the grade distributions below in Chapter 2, on the on historical patterns of grades and grading in First-Year Composition at UTK in the 1920s and 30s. I share the original tables here, to support and further illustrate my argument that grades typically fall along a bell curve and are skewed towards the lower end of grades, amplifying the narrative that students are unable to reach the standards established by the university for good writing.

Table G.1 Grade distributions, AY 1928 – 1929

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111 (39 st's)	2	4	13	9	0	6	5
Perct 111	.051	.103	.333	.231	0	.154	.128
ENGL 112	0	5	17	12	0	3	3
ENGL 113	1	8	15	5	3	9	3
Total	3	17	45	26	3	12	11
Perct	.026	.145	.385	.222	.026	.103	.094

117 students total.

Table G.2 Grade distributions, AY 1929-1930

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	4	9	8	1	0	0	0
Perct 111 (22 st's)	.182	.409	.364	.045	0	0	0
ENGL 112	3	7	10	0	0	1	1
ENGL 113	3	5	7	5	0	0	0
Total	10	21	25	6	0	1	1
Perct	.156	.328	.390	.094	0	.016	.016

64 students total. Hodges only taught 1 section each of 111, 112, and 113 this AY. This chart does not include three students whose names are listed at the bottom of the 111 sheet and about whom no other information is provided

Table G.3 Grade distributions, AY 1930 – 1931

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	0	10	21	8	0	2	1
Perct 111 (42 st's)	0	.238	.5	.190	0	.048	.024
ENGL 112	2	9	18	7	1	0	2
ENGL 113	3	8	10	12	2	0	1
Total	5	27	49	27	3	2	4
Perct	.043	.231	.419	.231	.026	.017	.034

117 students total.

Table G.4 Grade distributions, AY 1931 – 1932

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	2	6	22	6	0	0	0
Perct 111 (36 st's)	.056	.167	.611	.167	0	0	0
ENGL 112	2	8	19	6	0	1	1
ENGL 113	2	5	15	8	2	3	1
Total	6	19	56	20	2	4	2
Perct	.055	.174	.514	.183	.018	.037	.018

109 students total.

Table G.5 Grade distributions, AY 1932-1933

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	2	2	10	4	3	4	4
Perct 111 (29 st's)	.069	.069	.345	.138	.103	.138	.138
ENGL 112	0	1	1	4	0	0	3
ENGL 113	2	2	13	11	1	2	2
Total	4	5	24	19	4	6	9
Perct	.056	.07	.338	.268	.056	.085	.127

71 students total. Hodges only taught 1 section of 112 this year, and to ensure that the randomized student was present in all three sections, I limited the randomizer to the number of students enrolled in the section of 112 that he actually taught, then checked that student was indeed enrolled across the other sections.

Table G.6 Grade distributions, AY 1933-1934

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	1	8	20	9	0	4	2
Perct 111 (44 st's)	.023	.182	.455	.205	0	.091	.046
ENGL 112	1	9	20	9	1	1	2
ENGL 113	1	8	20	17	1	7	2
Total	3	25	60	35	2	12	6
Perct	.021	.175	.420	.245	.014	.084	.042

143 students total.

Table G.7 Grade distributions, AY 1934 – 1935

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	1	5	18	9	1	8	2
Perct 111 (44 st's)	.023	.114	.41	.205	.023	.182	.045
ENGL 112	3	10	22	6	0	0	2
ENGL 113	5	8	20	9	0	2	0
Total	9	23	60	21	1	10	4
Perct	.070	.180	.469	.164	.008	.078	.031

128 students total.

Table G.8 Grade distributions, AY 1935 – 1936

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	2	9	13	9	0	7	4
Perct (Perct 111)	.045	.205	.295	.205	0	.159	.091

44 students, total. Hodges did not teach 112 or 113 this year

Table G.9 Grade distributions, AY 1936-1937

	As	Bs	Cs	Ds	Es	Fs	Other
ENGL 111	2	5	22	3	0	7	2
Perct 111 (41 st's)	.049	.122	.537	.073	0	.171	.049
ENGL 112	3	5	18	10	0	3	1
ENGL 113	2	6	12	14	0	2	4
Total	7	16	52	27	0	12	7
Perct	.058	.132	.430	.223	0	.099	.058

121 students total.

Appendix H: Codebook

This appendix includes a codebook for the primary codes used in this project — perceived affordances of ungrading and instructor values. I coded the data at other dimensions, such as the invention/development of a particular ungrading method, but as I do not use those codes in this project, I have omitted them from the appendix here.

Dimension 1: Affordances

Using the word "affordances," I aim to capture the particular opportunities or ways of being that ungrading opens up to teachers. As I've discovered, there's a lot of overlap with "values," but this dimension is less about the beliefs that instructors bring to the course and more about what they anticipate ungrading doing and/or what it actually does in practice for them, their students, and their course. The list below includes the code (in bold), then the definition after it.

- **Attends for social inequities outside of race.** *Ungrading* acknowledges and/or removes social inequities from the assessment process, so that students are assessed only on relevant work for the course. Items may be coded here even if they do not mention ungrading explicitly, so long as participants speak directly about the relationship between social inequities and grades/grading.
- **Attends for social inequities — Covid.** *Ungrading* acknowledges and accommodates the challenges that emerged or were amplified b/c of Covid, and that students or teachers faced. Covid may be glossed as "2020," "these times," etc. Items may be coded here that do not mention ungrading, so long as participants explicitly recognize the Covid-related challenges students face.
- **Develops student-teacher relationship.** *Ungrading* fosters a non-hierarchical relationship between students / teachers, minimizing teachers' work as arbiter in favor of work as coach / mentor / guide. (Please note that this code is sometimes numbered under "interacts with pedagogy," as a specific aspect or feature of pedagogy.)
- **Surfaces antiracist teaching practices.** *Ungrading* cultivates opportunities for students with diverse raciolinguistic backgrounds to succeed and thrive. Teachers have the opportunity to pursue antiracist strategies explicitly.
- **Interacts with pedagogy.** *Ungrading* supports those ways of being in the classroom, both students' ways of being and teachers' ways of being, that facilitate learning and writerly growth, e.g. lessened anxiety, risk-taking, a focus on transfer, etc. Items may be coded here if instructors discuss how ungrading constrains or limits student learning.
- **Values labor (instructor labor).** *Ungrading* makes it possible for teachers to spend their time and energy how they want, and/or to develop into the kind of teacher they want to be. Both quantitative (the amount of time) and qualitative (what goals or tasks the time is spent on). Discussion of how ungrading allows teachers to give feedback instead of grade/justify a grade is coded here.
- **Values labor (Student labor).** *Ungrading* makes the course easier to pass, often via assessment design that removes barriers and allows student work to pay off in the form of success. Items may be coded here if instructors fear that ungrading undermined success.

Dimension 2: Values and Ways of Being

In developing this dimension, I relied on language such as "I wanted," "I loved" or "I am pleased with" to indicate strong emotional feelings associated with values and beliefs, or valuative words such as "really personalize feedback." At the same time, I chose to overcode (rather than undercode) along this dimension, pulling in segments which do not necessarily use this language. This choice is a practical one, in that it better recognizes how language works. The affordances or opportunities created by ungrading generally hinge on some assumed value, which does not necessarily need to be explicitly stated. When, for instance, an instructor says that ungrading makes them less stressed, the desire for self-care is implicit and does not need to be spelled out. Nobody values being *more* stressed. If there seem to be competing values within a given segment, I chose the dominant or central one (though I think that double- or even triple-coding the values, to determine where they cluster, could provide important additional analysis). Finally, some segments are coded at particular values less because of what *that particular segment* reveals but because of the larger, conversational context; I used my knowledge of what comes before or after each segment to guide my coding choices.

Also, while most data coded in this dimension includes *positive* descriptions, e.g. people describe their goal or value in positive terms, at least a few items are coded because a *negative* description reveals the person's value. Carter's description, for instance, of having their teaching choices constricted through their work with a supervisor indicates that Carter values agency, even though that value is not currently realized.

As with the list above, the code is listed in bold, with the description or definition following.

- Prioritizes fairness. *Instructors value education* that avoids penalizing students and/or finds alternative ways to acknowledge student accomplishments outside of quality measurements. The teacher is no longer the judge or arbiter of quality. (Elliot, 2016).
- Recognizes / values students' subjectivity. *Instructors value education that* acknowledges the place(s) students come from and their individual needs over objective, one-size-fits-all institutional expectations. Teachers have an obligation to fill the gap.
- Comes alongside students as a guide/mentor rather than as a judge. *Instructors value education that* situates them in a non-hierarchical position relative to students, a mentor/coach or "guide on the side" versus an arbiter of grades. Instructors value being open and vulnerable with students.
- Practices self-care. *Instructors value education that* acknowledges the stressors and mental health needs that students and teachers bring into the classroom with them and seeks to meet them if possible.
- Empowers students. *Instructors value education that* ensures students have the knowledge and space/structure to make choices that let them grow as writers and reach their goals. Antiracism is coded here.
- Centers rhetorically-informed writing education. *Instructors value education that* centers best practices and goals from rhetorical or writing studies scholarship, such as transfer, genre, reflection, and the distinction between higher and lower order concerns.
- Practices sound pedagogy. *Instructors value education that* aligns with best practices in education. They value being effective teachers and receiving high-quality student work.

This code encompasses teaching practices that "translate" across disciplines, e.g. it is not rhetorically specific.

- Enjoys teacher agency. *Instructors value* the freedom to make decisions for themselves about how they do their work and/or how they spend their time. This extends to environmental factors (they value support from students or colleagues) and personal ones (they value feeling confident in their own decisions).

Vita

Megan Von Bergen received her MA degree in English Literature from Kansas State University, whe she wrote her thesis on T.S. Eliot's *Four Quartets.* After graduating, she taught first-year composition and introductory literature courses as the sole English faculty at a faith-based private college in northeast Iowa. This work sparked her interest in rhetoric and composition as a discipline, and she returned to pursue a doctoral program — at first with a focus on religious rhetorics, and then with a focus on learning-centered, equitable forms of writing assessment. Broadly, Von Bergen is interested in questions about how to generate and sustain institutional change to support diversity and inclusion and enable students, teachers, and administrators to thrive. She looks forward to continuing this work in her new position. Outside of her academic work, Von Bergen enjoys trail running and reading science fiction.