

Scaling Responsible Data Science Education: The Role of a Teaching Assistant in Bridging the Sociotechnical Divide

Abigail Brooks-Ramirez, University of California, Berkeley

Abigail Brooks-Ramirez is an Undergraduate student at the University of California, Berkeley completing her B.A. in Computer Science. Her research interests, strongly informed by her time as both a student and Undergraduate Student Instructor, center on inclusive pedagogy and creating tools for social impact.

Rebecca Dang, University of California, Berkeley

Rebecca Dang is an undergraduate student studying Electrical Engineering and Computer Science at the University of California, Berkeley. Her research interest is computing and data science education, including building tools that enable teaching at scale and improving pedagogical training for teaching assistants. She is advised by Professor Lisa Yan.

Bryan Adolfo Ventura Benitez, University of California, Berkeley

Bryan Ventura is a 4th-year undergraduate Computer Science student at the University of California, Berkeley, advised by Professor Lisa Yan. His research interests include CS education, AI, and large language models, with a focus on creating inclusive and equitable learning environments in computing education.

Lisa Yan, University of California, Berkeley

Scaling Responsible Data Science Education: The Role of a Teaching Assistant in Bridging the Sociotechnical Divide

Abstract

Students in undergraduate-level data science (DS) programs undergo highly technical engineering education only to enter the workforce underprepared to participate in technological development inherently enmeshed with social contexts. Responsible data science curriculum seeks to bridge this skill gap by directly teaching ethical, accountable, and socially responsible DS practices alongside technical learning objectives, often within the same course. However, in undergraduate programs with hundreds of students per course, much of a student's learning happens outside of any interactions with faculty instructors. In particular, graduate or undergraduate teaching assistants (TAs) may facilitate much of small group classroom education: designing and shaping classroom activities, setting evaluative standards, and establishing classroom culture. The practice of teaching responsible data science curriculum in large-scale undergraduate DS programs therefore necessitates recruiting TAs fluent in teaching across the sociotechnical divide—to design activities and build learning spaces that meaningfully practice learning objectives situated across multiple fields of study.

This paper examines how TAs teach and experience responsible data science curriculum within a DS program at a North American public R1 institution. In particular, we seek to investigate how TAs acquire and subsequently implement sociotechnical fluency in their pedagogy, and what this might mean for implementing sociotechnical engineering education at scale. The paper contextualizes its study within a new TA professional development course, designed for first-time TAs to practice skills for teaching technical curriculum alongside those for developing inclusive classrooms. The course itself is an implementation of responsible data science curriculum; it leverages culturally relevant pedagogy and growth-oriented educational frameworks to deliver both practical pedagogical training and reflective reading and writing activities. Through this course, TAs reflect on their roles to both foster student technical understanding and create a broader social space for engineering education.

Introduction

As the artifacts and processes of engineering disciplines become more deeply entrenched in society, the professional obligations of an engineer become more complex and intersectional. Subsequently, in academic engineering settings, the discussion of “engineering ethics” within undergraduate education has morphed into a broader goal for *sociotechnical education*, which informs students of the collective social responsibilities held by engineers and the ways in which an engineer’s discipline shapes and is shaped by society [1]. In the context of data science (DS) undergraduate programs, *responsible data science curriculum* integrates engineering ethics and sociotechnical education by embedding ethical and socially responsible DS practices alongside technical learning objectives, often within the same course.

In DS undergraduate programs, this sociotechnical education is often offered as a standalone curriculum addressing human contexts within the discipline or as an integration into existing technical coursework [2]. While the former—in its most ideal—can involve solo design and instruction from experts in fields like science, technology, and society, the latter must be co-designed and co-owned by the technical instructor. The instructor must therefore also have a base level of *sociotechnical fluency*, which we define as the ability to recognize and integrate the interdependence of technical and social systems in a given discipline. This need for sociotechnical fluency among technical educators is exacerbated when we expand the definition of an instructor from the faculty member—who intellectually leads the course design—to include *teaching assistants* (TAs), who design their own smaller classroom spaces for close interaction with individual students.

Given the breadth of the term “sociotechnical,” we delineate the term into internal and external sociotechnical systems. *External* sociotechnical systems refer to human stakes within the larger engineering sphere, spanning industry and research, which are commonly included in discourse surrounding engineering ethics [3]. *Internal* sociotechnical systems refer to individual relations within the engineering classroom which are informed by social contexts and are often enforced by the students and TAs themselves. Consequently, internal sociotechnical systems may reinforce patterns of structural inequality within the engineering classroom [4]. Related, educators teaching at this intersection generally understand their responsibilities along two threads: *what* content is taught—thereby connecting course content to the external sociotechnical systems as related to the specific discipline—and *how* this content is delivered—thereby realizing the internal sociotechnical systems of the specific classroom [5]. In large technical courses with longstanding curricula, TAs may provide minimal input on course content, but they play a large role in teaching the content. In other words, through fostering supportive internal sociotechnical spaces, TAs can shape the coproduction of knowledge and sociotechnical systems, influencing how students engage with both internal and external sociotechnical systems through their behavior.

In this paper, we examine sociotechnical fluency among TAs in a large DS undergraduate program at an R1 North American institution. This paper considers a TA-level intervention that addresses *how* sociotechnical education is taught through focusing on the internal sociotechnical dynamics (**Figure 1**). We study a professional pedagogy training course which leverages community-centered learning amongst data science TAs and trains the TAs to create similar

internal dynamics within their own classrooms. This is achieved through several means: developing a *community of practice* among TAs, an educational theory that describes collective learning [6], developing awareness of the TAs’ positionality in the broader degree program, , and guiding TAs in designing and managing the internal sociotechnical systems within their own classrooms such that the environment is conducive to responsible data science education.

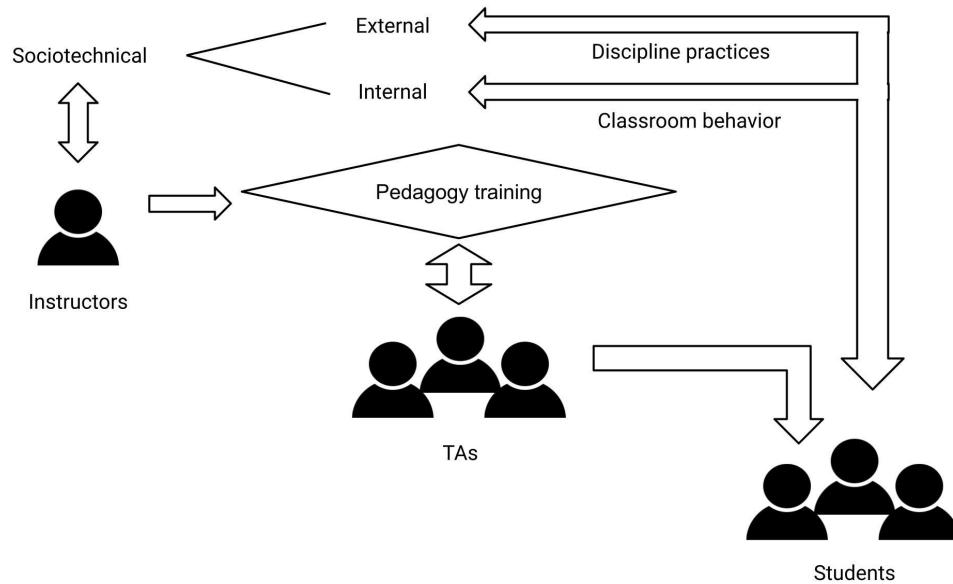


Figure 1: Pedagogy and sociotechnical systems. *A pedagogy training course serves as a TA-level pedagogical intervention. Students are influenced by and go on to influence internal and external sociotechnical systems.*

To understand the formation and implementation of effective sociotechnical pedagogy, we studied DS TAs for two semesters: the semester they were enrolled in the pedagogy training course and the subsequent semester they taught [7]. We analyzed written assignments from TAs throughout the course, ethnographic notes from our observations of TAs’ classrooms, and interviews with TAs to explore the following research questions:

RQ1. How do TAs acquire sociotechnical fluency in data science?

RQ2. How do TAs then adapt their pedagogy to incorporate this understanding?

Our results provide recommendations for effectively developing internal sociotechnical systems with TAs. We then briefly discuss the role of TAs developing content to further connect courses to external sociotechnical systems. In sharing these results, we hope to provide guidance for other engineering disciplines, offering a scalable model for creating space for sociotechnical learning in engineering education.

Related Work

Often, “engineering ethics” is conflated with sociotechnical education. In prior work, higher education institutions have taken several approaches to implementing this interpretation of sociotechnical education, such as standalone courses within and outside of engineering departments, embedded material within an engineering course, and courses taught by a team of

professors across technical and non-technical disciplines [5]. Due to limited prior work on responsible data science curricula, we instead examine the related field of computing ethics. Being a subfield of engineering ethics, computing ethics shares educational methods like standalone courses, embedded ethics [8], [9], and team teaching [10]. The latter two methods are intended to combat the notion that ethics, traditionally considered a humanities subject, is separable from engineering practices.

In teaching the intersectionality of engineering disciplines, educators often aim to teach the intersection between science and society, leading educators towards sociotechnical education. To understand the challenges of implementing sociotechnical education we draw upon a layered social organization of educational processes utilized in education research [4]. The macro-level describes societal assumptions and structural processes which are often reinforced by faculty and institutional structures. The micro-level describes student and individual experiences. At a macro level in engineering ethics, a lack of standardization for ethics exists across institutions, and as a result, faculty often face uncertainty in choosing how to integrate ethics into their programs [11]. Engineering faculty, often unfamiliar with ethical theory, further face uncertainty in what content to teach and designing grading schemes, whose subjective nature differs from traditional objective grading systems in engineering [11]. Consequently, these views are replicated at the micro-level and students may also perceive STEM fields as inherently “neutral” and “objective” rather than sociotechnical. Students understand their education simply as a means to acquire a job rather than a vehicle through which they can improve society, further widening the gap between liberal and engineering education [12], [13], [14].

The meso level, which bridges the macro and micro levels, has the potential to disrupt assumptions in both macro and micro levels. TAs, who naturally act as a bridge between institutional structures and students, are a critical meso-level intervention in sociotechnical education that we focus on in this study. Prior work has shown that TAs often improve student learning outcomes, engagement, and offer personalized support [15], [16]. At some undergraduate computing institutions, TAs who use inclusive teaching practices improve retention in STEM courses, especially for underrepresented groups [4], [15]. With a more approachable power dynamic compared to professor-to-student, TAs are also uniquely positioned to foster ethical debate [12], yet there is not much existing literature exploring the possibility of utilizing TAs to improve sociotechnical education.

One study found that engineering ethics TAs often struggle with whether to stay “neutral” in class discussions [17]. Identifying the need to support computing ethics TAs as class sizes grow, another study developed an online community of TAs across institutions on Discord, allowing them to share resources and ideas [18]. Another study interviewed computing ethics TAs across two R1 universities, finding that even TAs within the same institution received varying levels of training (with some TAs not receiving any training at all), echoing existing literature discussing TAs’ limited preparation to teach ethics [19]. While these papers point out that TAs teaching sociotechnical concepts could benefit from training and community, the literature does not investigate this further. Additionally, the literature focuses only on TAs for standalone engineering and computing ethics courses, while our research supports the idea that TAs can play a pivotal role in bridging the sociotechnical divide — whether they directly teach computing ethics or teach a technical course.

Moreover, TAs often play a multifaceted role, with various responsibilities that are often institution-specific. Regarding contributions to course content, there is minimal prior work on the extent to which data science TAs contribute to the development course materials (much less sociotechnical course materials). A systematic review of studies involving computer science TAs found that some were responsible for content development, [15], but data science curriculum development is often an institutional or faculty responsibility [20], [21], [22].

Institutional Description

We study UC Berkeley’s DS undergraduate program, a large, interdisciplinary program that offers longstanding and foundational courses co-designed by experts within the field. Our public R1 North American institution supports over 30,000 undergraduate students campuswide and annually awards over 890 undergraduate computer science and over 840 undergraduate data science degrees [23]. The DS degree is institutionally defined as an undergraduate training in computational and inferential thinking—combining computing and statistics—in addition to an individual domain emphasis of choice (e.g., business and industrial analytics, cognition, computational biology, etc.) [23]. Degree-seeking students are required to engage with the human contexts and ethics of data (HCE) through a standalone HCE course; several foundational and elective DS courses offer course modules embedded with responsible data science practices and sociotechnical contexts [3]. Combined, these components aim to prepare students for the social and technical contexts they may encounter in their future work. At present, DS is a primarily undergraduate program at our institution.

Each semester, DS first-year and third-year courses enroll approximately 1,000 and 500 students, respectively, necessitating the use of TAs to execute much of the student-level teaching and management. Over 80% of DS TAs are senior undergraduate students who have previously taken the course in question. However, many of these TAs have not encountered sociotechnical work in an academic setting (e.g., HCE) and therefore are ill-equipped to teach sociotechnical curriculum. Below, we discuss the professional pedagogy course Data 375 as a space for equipping TAs with the pedagogical skills required to manage internal sociotechnical systems and thereby space for intersectional learning and discussion.

Pedagogy Training Course Description

In large-scale DS courses, lab and discussion sections are often facilitated by undergraduate or graduate TAs. At our institution, all TAs must take a professional pedagogy training course before or during their first semester teaching. Prior to Spring 2024, data science TAs would take their professional training course in the Computer Science or Statistics departments. The pedagogy course studied in this paper, Data 375, is a data science-focused pedagogy course offered every semester that was first designed and piloted in Spring 2024 [7], [24]. Data 375 teaches educational theory and professional development while introducing frameworks for social justice and inclusion. Beyond technical pedagogy, learning goals for TAs also include understanding internal sociotechnical systems within DS classrooms. It incorporates readings on inclusive educational theories (e.g., culturally relevant pedagogy [25]) and reflective assignments (e.g., journal entries and curriculum development projects). By integrating these reflective

practices, TAs explore sociotechnical systems within the classroom and address the sociotechnical gap by acquiring the knowledge and skills necessary to teach responsible DS education. The course includes a variety of assignments that complement in-person discussion; in this study, we focus on responses to surveys and weekly journals.

Researcher Positionality

As researchers, we acknowledge that our positionality, shaped by our backgrounds, experiences, and institutional roles, directly informs our approach to this work. This research is conducted by a team of scholars from the Electrical Engineering and Computer Science department at the institution studied. The three undergraduate researchers in this study identify as members of historically underrepresented groups in engineering, bringing with them life experiences that shape their understanding of equity, access, and pedagogy in DS education. Two of these undergraduate researchers also currently serve as TAs within the DS degree programs, further positioning them at the intersection of both institutional structures and student experiences. Their dual role as both educators and students enables a critical perspective on how teaching assistants navigate the sociotechnical divide, particularly in how they balance technical instruction with fostering inclusive learning environments. The faculty researcher in this project also identifies as a member of a historically underrepresented group in engineering faculty; she regularly teaches large computing and DS courses on campus and also designed and taught the pilot pedagogy course mentioned in this study [7].

Methodology

To examine sociotechnical fluency among teaching assistants (TAs) in our institution, we conducted two studies: first, the initial study in Spring 2024 which analyzed the pilot offering of Data 375 as a meso-level intervention, and second, the follow-up study in Fall 2024 which evaluated the long-term effects of the intervention. **Table 1** shows the demographic breakdown of study participants. **Table 2** describes in detail the data collected in each phase, including assignments, surveys, and journals.

Initial Study (Intervention): In Spring 2024, we collected data from weekly journal entries and surveys as assigned by the course. TAs completed entry and exit surveys via Google Forms at the beginning and end of the semester. Both surveys included four mandatory open-ended questions designed to explore TAs' personal and professional backgrounds and their pedagogical practices, along with an optional demographic self-report. Throughout the semester, TAs were required to answer journal prompts, which asked TAs to explore teaching goals, teaching priorities, and reflections on their experiences in academic settings from a teaching and learning perspective. In addition to this data, we also collected student assignments for the purpose of analyzing how students were engaging with the content covered in the course.

Follow-Up Study: In the Fall 2024 semester, we followed up with TAs who participated in the Spring 2024 data collection process and were currently teaching, as well as any current DS TAs who may not have participated in the Spring 2024 study. Of the 14 TAs who participated in the Fall 2024 study, 4 had taken Data 375 during Spring 2024. For this paper, we performed a qualitative analysis using the data from both our initial and follow-up studies.

In the follow-up study, TAs filled out an interest survey, where TAs could opt-in to participate in classroom observations and/or interviews. This part of the study was aimed at assessing the enduring impacts of their training on their teaching approaches and effectiveness, providing valuable insights into the long-term benefits and areas for improvement in our pedagogical training intervention through Data 375. This study was approved by our Institutional Review Board; all data collected, qualitatively analyzed, and presented in this study has been de-identified to safeguard the privacy of the participating TAs.

Characteristic	Category	Count	
		Spring 2024	Fall 2024
Department of Course Taught	Data Science	13	9
	Other	3	5
Degree Status ¹	Undergraduate	14	13
	Graduate	2	1
Gender	Gender minority ²	9	9
	Not gender minority	7	5
Major ³	Data Science	12	8
	Computer Science	4	8
	Other	9	6
Number of years at institution	3 years	9	4
	4 years	4	7
	Other	3	3
Race	Underrepresented racial minority (URM) ⁴	3	N/A ⁴
	Not URM	13	N/A ⁴

1. Degree status refers to the current level of enrollment at the time of data collection.
2. This study defines gender minorities as female-identifying, non-binary, or LGBTQ+. For the Fall 2024 study, we used pronouns as a proxy for gender.
3. Many students majored in 2 or more fields. In this case, we double counted.
4. This study defines underrepresented racial minorities as African American, Chicano/Latino, Native American/Alaska Native, and Pacific Islander. Race was not collected in the follow-up study.

Table 1: Participant Demographics. *Most students who took Data 375 were undergraduate TAs majoring in computer and/or data science who were teaching undergraduate data science courses.*

Spring 2024: Initial Study	
Data Type (sample size)	Notes
Entry survey responses (16)	- Required survey for all Data 375 students at the start of the semester. - In this survey, 16 out of 20 students in Data 375 consented to participate in our initial study. - Included 4 mandatory written responses to questions asking about teaching priorities and inclusivity.
Exit survey responses (16)	- Required survey for all Data 375 students at the end of the semester. - Included the same 4 mandatory written response questions as the entry survey.
Journal entries (16)	There were 13 journal entry prompts per student over the semester.
Sociotechnical Project submissions (6)	- In this assignment, Data 375 students redesigned part of a discussion worksheet from the class they teach to embed sociotechnical content. - Students in Data 375 completed the project in 10 groups of 2. Of these, 6 groups had both group members consent to the initial study.
Fall 2024: Follow-up Study	
Data Type (sample size)	Notes
Entry survey responses (14)	- Voluntary survey sent to current computer and data science TAs in the middle of the semester. - In this survey, participants gave consent to a classroom observation, interview, both, or neither. - To incentivize participation, participants who consented to a classroom observation and/or interview received a small gift card.
Exit survey responses (5)	- Voluntary survey sent to current computer and data science TAs at the end of the semester. - To incentivize participation, participants who completed the form received a small gift card.
Classroom observation notes (5)	- Classroom observations were conducted in-person in the last half of the semester. - One author observed each TA and took ethnographic notes during and after the observation.
Interview transcripts and notes (8)	Interviews were conducted virtually over Zoom by the authors. The Zoom automatic transcription feature was used to generate transcripts, and separate notes were also taken by each interviewer.

Table 2: Data collection. *Data collected for qualitative analysis in the initial study in Spring 2024 and the follow-up study in Fall 2024.*

Findings

RQ1: How do TAs acquire sociotechnical fluency in data science?

The development of sociotechnical understanding is crucial for TAs, as it enables them to effectively guide students in mastering both the technical intricacies and the broader social implications of various engineering disciplines, including DS. TA interviews and Data 375 assignments suggest that TAs' lived and educational experiences provided them with varying levels of sociotechnical fluency prior to even entering the Data 375 classroom.

Identity Shapes Teaching Practices: TAs who held marginalized identities (see **Table 1**) frequently drew upon their experiences with systemic inequalities to inform their teaching goals and practices. For instance, one TA in the welcome survey for Data 375 shared, "My identity has significantly shaped my academic and professional paths. Coming from an underrepresented community, I've always felt the need to advocate for more inclusivity within academic settings." Another TA, during an interview, discussed their experiences in the technology industry as an underrepresented racial minority, discussing how their positionality served as a lens to observe and realize inequality within technology practice. Industry experiences proved formative for this TA, and they discussed their desire to increase technical literacy in underserved communities in their current and future work. TA experiences illustrate the role of positionality and lived experiences in building familiarity with both internal and external sociotechnical systems, suggesting that sociotechnical fluency may be acquired outside of the classroom, often by those who confront systems of inequality in their daily lives. In particular, those who confront systemic inequalities often develop a keen awareness of classroom power dynamics and the importance of inclusive teaching strategies, even before formal pedagogical training.

While these experiences provide a foundational perspective and are a common sentiment amongst students in the course, TAs often find themselves lacking the precise vocabulary, structured approaches, and frameworks necessary to effectively translate and implement these insights into their educational settings and pedagogy. As an internal pedagogical intervention, Data 375 addresses this gap by immersing TAs in a curriculum that underscores culturally relevant pedagogy and formative assessments, strategically designed to highlight the sociotechnical divide in DS. This curriculum includes readings such as *Anti-racist pedagogy: from faculty's self-reflection to organizing within and beyond the classroom* [26] and *Data Science Ethos Lifecycle: Interplay of Ethical Thinking and Data Science Practice* [27] which provide theoretical grounding, while structured reflections and peer discussion reinforce the practical application of sociotechnical principles in teaching.

Communities of Practice: In the case of Data 375, TAs are brought together through their shared interest in teaching and immersed in collective learning through Data 375's course structure, which strongly emphasizes collaborative learning, both synchronously through in-class discussions, peer observations, and asynchronously via QQCRs (Question, Quote, Comment, and Reply exercises). This approach ensures that TAs can learn from their peers' diverse experiences, backgrounds, and insights, thereby enhancing their ability to apply these teaching practices effectively and fostering a reflective practice amongst TAs that is essential for developing sociotechnical fluency. TAs reacted positively to this model, with several mentioning in

interviews and surveys that they benefited from forming a teaching community and hearing from fellow TAs' lived experiences. One TA on the exit survey reflected on how their approach to teaching was shaped by "sharing some of [their] experiences [with other TAs] and pushing students to advocate for themselves and their learning."

Similarly, another TA from the Fall 2024 cohort highlighted the profound impact of Data 375 and its community-based learning:

After taking Data 375, I realized the importance of integrating culturally responsive teaching into my approach... This came into sharp focus when a peer shared her experience in a biology lab, where she felt overlooked despite the evident disparities in student understanding... It underscored the need for workshops focused on culturally responsive practices, not only addressing race and gender but also the unique cultural context of each classroom.

The TA's commitment to changing the teaching culture at the institutional level exemplifies the transformative potential of the pedagogical skills developed in Data 375 in approaching internal sociotechnical systems.

RQ2: How do TAs then adapt their pedagogical practices to incorporate their sociotechnical fluency?

In teaching responsible data science effectively, TAs must balance multiple layers of sociotechnical integrations. In this section, we focus on TAs' internal sociotechnical fluency and how it manifests in their classrooms. Approaching internal sociotechnical systems requires TAs to acknowledge systemic barriers faced by their students in DS classrooms and the ways in which classroom dynamics may reinforce these barriers. Existing research acknowledges the existence of barriers in DS education for students from underrepresented backgrounds, which include holding a lower sense of belonging, facing imposter syndrome, and social exclusion within the field [4]. In approaching these challenges, several TAs in our study leveraged culturally relevant pedagogy—as informed by Data 375—to approach inclusion on a student level. Throughout classroom observations and Data 375 responses from TAs, inclusive pedagogical practices branched into three main categories.

Denouncing Status Hierarchy: "Status hierarchy" is the process, enforced by institutional systems, that frames instructors and TAs as possessors of expert knowledge, subsequently positioning students as passive recipients of that knowledge, rather than collaborators within the educational process [4]. As students participate in this hierarchical classroom dynamic, the power relations practiced in the classroom become a part of student identity, naturalizing a subversive role for students within the classroom and limiting the agency they have within their learning [28]. During the formation of their pedagogy, TAs reflected on their positionality and power as an instructor. In practice, TAs worked to remove perceived barriers to TA accessibility and frame themselves as a resource for support, rather than a figure of total authority, effectively empowering students in their learning practices.

TAs first expressed awareness of their positionality as instructors, and the power that came with their role, in early Data 375 discussions. Responding to a passage from *Teaching Critical Thinking* by bell hooks [29], one student commented, "if we're teaching students who are at the

same academic level as ourselves, we don't hold that much academic authority over them. It feels a lot less burdensome to think of yourself as a bouncing board for students' ideas than to think of yourself as an authoritative judge where all ideas come to an end." Many of the TAs in our study were undergraduate and master's students and drew from their experiences as students to inform their pedagogy goals. TAs often found it easier and more beneficial to pose themselves as a co-constructor of knowledge rather than a figure of authority.

In the same thread, another TA mentioned during their interview that forming strong connections with students is often missing from the traditional collegiate experience, a gap they felt strongly during their undergraduate experience as a first-generation college student. In an effort to close this gap in their own classroom, this TA arrives at their section early to initiate informal conversations with their students, inquiring about their weekends and interests to understand their students outside of the educational sphere, challenging the monolithic view of students that is often seen in education [4]. In turn, this TA is able to gauge their students' state of mind prior to beginning lecture and meet the specific needs of their students. Through engaging the classroom in this manner, students are understood as individuals, as is the TA, empowering students to collaborate with their TA in the learning process.

Creating Safe Classrooms: We derive the term "safe classrooms" from prior work in third spaces and safe spaces. Third spaces, generally defined as social spheres separate from work and home where community needs are met, have been studied within academic settings and understood as classrooms that welcome intersectional engagement and dialogue [28]. Safe spaces originate from Queer communities in order to co-construct spaces that meet the mental and physical needs of members. Bridging these spaces, we define safe classrooms as educational spaces in which individuals are able to engage authentically with educational content and with their peers. Prior work has examined the importance of dialogue in constructing safety in the classroom, as student and instructor contributions to classroom learning have the potential to reinforce feelings of belonging in the classroom [28].

Throughout pedagogy training, TAs reflected on the importance of ensuring students felt understood within the classroom. In one reading response, one TA commented "In order to have the greatest impact on our students, we should be doing our best to get to know them and contextualize their learning within their lived experiences." Culturally relevant pedagogy was a crucial component of TA understanding of internal sociotechnical processes that inform classroom culture. Understanding the contexts that shape student learning allowed TAs to reflect on internal systems within the classroom and implement teaching practices to challenge historically marginalizing systems within the classroom.

In practice, TAs implemented pedagogical strategies to lower student anxiety, understanding that learning data science was often a high-stress activity for students. Several TAs began sections with a "temperature check" in order to determine which topics students needed support with. Many of these temperature checks were non-verbal (such as asking students for a "thumbs up" or "thumbs down") to lower barriers to participation and enable participation from students who may be hesitant to speak up. TAs further encouraged questions from their students, positively affirming questions and acknowledging that difficulties in understanding course content are common. Through co-constructing safe classrooms with their students, TAs create spaces that

lower anxiety to participation and encourage community formation, countering systemic barriers in the classroom.

Communities of Practice: We further explore communities of practice, discussed in RQ1, in the context of the TAs' own classrooms. Through naturalizing collective learning in the classroom, students may engage with content more thoroughly while establishing learning communities, which are critical to student retention [4]. Communities of practice was a core component of Data 375, and TAs positive experiences with learning in community often drove them to implement similar communities in their own classrooms.

TAs who had completed Data 375 structured their classrooms to emphasize collaboration, implementing techniques that had been modeled in Data 375. Several TAs utilized a “think, pair, share” model, which allows students to work on a problem set individually, then discuss their work with a peer, and finally collaborate with another student pair, forming a group of four. One TA asked their students to brainstorm through a shared medium (i.e. Google Documents) in order to collectively work towards solving a problem set and ask clarifying questions. TA actions that show care for the classroom community strengthen students' learning experiences, and shaping the classroom into a community of practice is critical for TAs to teach effectively.

Reflecting on their pedagogy in a journal entry for Data 375, one TA wrote:

I think I have built a pretty strong relationship with the majority of my students. I know almost all of their names and I have not had silence when I asked a question for any of the last 3 or 4 discussions. I feel like especially today when I taught the ethics question students were really willing to open up and talk about their perspectives.

When students feel comfortable within the classroom, learning becomes more engaging and students are able to bring their full selves to the classroom. Through discussion and community-based learning, students may challenge normative assumptions about data science and experience an intersectional data science education.

Discussion

This paper's analysis has focused on how TAs have adopted the techniques they learned in Data 375 to shape the manner in which responsible data science is taught, thereby shaping *internal* sociotechnical systems. While this work qualitatively examined the experiences of TAs who took Data 375 and their subsequent teaching practices, our study is limited by its small sample size and our own institutional context. Our study methodology also solicited volunteers for interviews and classroom observations, thus potentially biasing towards highly-motivated TAs who were confident in their teaching abilities. We imagine that our large population of motivated, undergraduate TAs facilitating effective, near-peer teaching [15] is unique to our institution, and many engineering programs nationwide are supported primarily by graduate TAs or by faculty alone. Further constraints on replicating a TA-level intervention at other institutions are discussed in our prior work [7], which focuses more on the design of the professional pedagogy training course.

We now briefly explore how TAs might integrate *external* sociotechnical systems in their teaching, through the context of a Data 375 pilot assignment. As noted previously, while some TAs begin teaching with a general awareness of ethical issues through coursework or industry experiences, these engagements are often informal and do not have structured pedagogical strategies. The pilot offering of the pedagogy training course in Spring 2025 attempted to formalize this engagement by providing a structured assignment to design for ethical considerations within their discipline. In pairs, TAs were tasked with redesigning technical discussion worksheets from a DS course to embed HCE ideas alongside core technical content.

Through this assignment, TAs were not merely informed about the importance of sociotechnical fluency and ethics in computing, they were given a scaffolded, hands-on opportunity that enabled them to critically examine how technical and social systems intersect in both content and pedagogy. One strong assignment example redesigned a discussion worksheet question from the large, foundational introductory data science course [23]. Previously, the worksheet only introduced common visualization tools in Python, but now asks students to reflect on the ways technical choices, such as plot type or axes, may inform political narratives. The TAs who designed the worksheet posed reflective questions throughout the worksheet, directing students to consider data science as a rhetorical tool that is often dependent on certain socio-cultural contexts.

Despite the example shared above, assignment quality was mixed. Most TAs struggled with articulating HCE questions that were specific enough for students to discuss, and several proposed social contexts which suggested a surface-level understanding of the social factors at play. Upon reflection with an HCE colleague, we determined that the assignment learning objectives were too open-ended. Ultimately, TAs are hired for their expertise in the target course, and not their expertise in HCE fields.

Challenges of Developing New Sociotechnical Curriculum

While the sociotechnical projects from Data 375 demonstrate the possibility for TAs to integrate sociotechnical content into their courses, personal and institutional contexts ultimately dictate whether TAs have the capacity to do so. In large-scale courses where a team of TAs often oversee sections, host office hours, and manage course logistics, TAs often lack the time, power, or resources to do much else. Throughout interviews, TAs expressed fatigue as they worked to balance teaching with their own academic responsibilities. One TA in Data 375 commented that they only see their students once a week; in that single meeting, they are expected to go through a large list of course exercises and discussion questions. They expressed that inspiring their students or expressing individuality “sounds too fanciful for a [TA]” given the required teaching responsibilities that they must achieve. These reported experiences reflect existing literature even outside of the data science discipline, with a study finding that engineering TAs struggled to balance heavy workloads and other commitments as a student [30] and a systematic literature review of computer science TAs finding the breadth of responsibilities they can have, including leading sections, grading, supporting instructors during lecture, and hosting office hours [15].

Moreover, large introductory courses at our institution create silos of TA responsibilities, meaning that few TAs are involved in content development. Instead, most TAs manage the

complex course logistics required for supporting large numbers of students, and course materials are often passed between semesters with minimal modifications. Ultimately, while pedagogy training courses like Data 375 can train TAs in the skills needed to design and embed sociotechnical content within engineering education, such a movement requires institutionalized intervention through higher faculty- and instructor-level investment, such as more curricular collaboration between social and technical disciplines.

Conclusion and Future Work

This work explores the potential of TAs to bridge the sociotechnical divide in DS education. TAs can serve as an effective intervention to designing internal sociotechnical systems in engineering classrooms. Our findings leverage a space for providing sufficient pedagogical training for TAs so they may develop their existing understanding of sociotechnical gaps in engineering disciplines and acquire practical pedagogical practices to effectively communicate these concepts both within and beyond academic settings. While the Data 375 course curriculum provides direct benefits to TAs by enhancing and solidifying their sociotechnical fluency and ability to integrate ethical considerations into technical teaching, it is essential to recognize the broader implications of this course intervention. TAs often serve as the primary educational interface for students in their respective courses, especially in larger university settings like ours, where TAs are responsible for numerous students over multiple academic terms. Thus, the skills TAs acquire and refine in the course can potentially have a cascading effect on the students they teach. In other words, effectively trained TAs can significantly influence students' understanding, fluency, and appreciation of sociotechnical aspects in their disciplines—in this case, DS. This effect has the potential to foster a generation of students who are more attuned to the social contexts of their work across various engineering fields.

These early findings set the stage for future work that can assess the long-term impact of sociotechnical training on professional practice, pilot the integration of such training into other engineering disciplines, and explore broader sociotechnical curriculum development. Future longitudinal studies could explore how such pedagogy courses influence how TAs teach in different academic contexts. Additionally, many of our DS TAs still take computing and statistics pedagogy training courses in lieu of the DS pedagogy training course studied; we plan to explore downstream effects of different pedagogy courses by studying a broader range of DS TAs in the classroom.

Finally, ongoing and future work should address how undergraduate TAs and graduate researchers can collaborate with instructors to develop long-lasting, meaningful sociotechnical course curricula in data science and other disciplines [3]. Given the barriers to developing such curricula as identified in this paper, we hope to further explore which kinds of TAs are motivated to design course content, how to train TAs to teach and structure interdisciplinary classrooms, and who should drive development—higher education institutions, faculty instructors, TAs, or a combination of the three.

References

- [1] G. Ermer and S. VanderLeest, "Using Design Norms To Teach Engineering Ethics," presented at the 2002 Annual Conference, Jun. 2002, p. 7.1253.1-7.1253.10. Accessed: Jan. 09, 2025. [Online]. Available: <https://peer.asee.org/using-design-norms-to-teach-engineering-ethics>
- [2] N. Brown, B. Xie, E. Sarder, C. Fiesler, and E. S. Wiese, "Teaching Ethics in Computing: A Systematic Literature Review of ACM Computer Science Education Publications," *ACM Trans Comput Educ*, vol. 24, no. 1, p. 6:1-6:36, Jan. 2024, doi: 10.1145/3634685.
- [3] C. Carson, A. Edmundson, R. Sridharan, A. Strang, and L. Yan, "Sociotechnical Integration in data science Education," presented at the 2025 ASEE Annual Conference, Jun. 2025.
- [4] C. von Vacano *et al.*, "Critical Faculty and Peer Instructor Development: Core Components for Building Inclusive STEM Programs in Higher Education," *Front. Psychol.*, vol. 13, May 2022, doi: 10.3389/fpsyg.2022.754233.
- [5] J. Li and S. Fu, "A Systematic Approach to Engineering Ethics Education," *Sci. Eng. Ethics*, vol. 18, no. 2, pp. 339–349, Jun. 2012, doi: 10.1007/s11948-010-9249-8.
- [6] "Introduction to communities of practice - wenger-trayner." Accessed: Jan. 15, 2025. [Online]. Available: <https://www.wenger-trayner.com/introduction-to-communities-of-practice/>
- [7] L. Yan, "Teaching Our Teacher Assistants to Thrive: A Reflexive, Inclusive Approach to Scalable Undergraduate Education," in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1*, in SIGCSETS 2025. New York, NY, USA: Association for Computing Machinery, Feb. 2025, pp. 1281–1287. doi: 10.1145/3641554.3701950.
- [8] M. Skirpan, N. Beard, S. Bhaduri, C. Fiesler, and T. Yeh, "Ethics Education in Context: A Case Study of Novel Ethics Activities for the CS Classroom," in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, in SIGCSE '18. New York, NY, USA: Association for Computing Machinery, Feb. 2018, pp. 940–945. doi: 10.1145/3159450.3159573.
- [9] B. J. Grosz *et al.*, "Embedded EthICS: Integrating Ethics Broadly Across Computer Science Education," Aug. 16, 2018, *arXiv*: arXiv:1808.05686. doi: 10.48550/arXiv.1808.05686.
- [10] R. Reich, M. Sahami, J. M. Weinstein, and H. Cohen, "Teaching Computer Ethics: A Deeply Multidisciplinary Approach," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, in SIGCSE '20. New York, NY, USA: Association for Computing Machinery, Feb. 2020, pp. 296–302. doi: 10.1145/3328778.3366951.
- [11] D. A. Martin, E. Conlon, and B. Bowe, "A Multi-level Review of Engineering Ethics Education: Towards a Socio-technical Orientation of Engineering Education for Ethics," *Sci. Eng. Ethics*, vol. 27, no. 5, p. 60, Aug. 2021, doi: 10.1007/s11948-021-00333-6.
- [12] R. P. Aleman *et al.*, "Mind the Gap: Exploring the Perceived Gap Between Social and Technical Aspects of Engineering for Undergraduate Students," presented at the 2021 ASEE Virtual Annual Conference Content Access, Jul. 2021. Accessed: Jan. 09, 2025. [Online]. Available: <https://peer.asee.org/mind-the-gap-exploring-the-exploring-the-perceived-gap-between-social-and-technical-aspects-of-engineering-for-undergraduate-students>
- [13] D. Ozkan and C. Andrews, "Perspectives of Seven Minoritized Students in a First-Year Course Redesign toward Sociotechnical Engineering Education," presented at the 2022 ASEE Annual Conference & Exposition, Aug. 2022. Accessed: Jan. 09, 2025. [Online].

Available:

<https://peer.asee.org/perspectives-of-seven-minoritized-students-in-a-first-year-course-redesign-toward-sociotechnical-engineering-education>

- [14] S. Vakil and R. Ayers, "The racial politics of STEM education in the USA: interrogations and explorations," *Race Ethn. Educ.*, vol. 22, no. 4, pp. 449–458, Jul. 2019, doi: 10.1080/13613324.2019.1592831.
- [15] D. Mirza, P. T. Conrad, C. Lloyd, Z. Matni, and A. Gatin, "Undergraduate Teaching Assistants in Computer Science: A Systematic Literature Review," in *Proceedings of the 2019 ACM Conference on International Computing Education Research*, in ICER '19. New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 31–40. doi: 10.1145/3291279.3339422.
- [16] C. J. Felege, C. J. Hunter, and S. N. Ellis-Felege, "Personal Impacts of the Undergraduate Teaching Assistant Experience," *J. Scholarsh. Teach. Learn.*, vol. 22, no. 2, Art. no. 2, Jun. 2022, doi: 10.14434/josotl.v22i2.31306.
- [17] Y. Peng, M. Poliakoff, and L. Rosenberg, "The Role and Challenge of Teaching Assistants in Engineering Ethics Courses. | EBSCOhost." Accessed: Jan. 09, 2025. [Online]. Available: <https://openurl.ebsco.com/contentitem/doi:10.5840%2Ftej2024327140?sid=ebsco:plink:crawler&id=ebsco:doi:10.5840%2Ftej2024327140>
- [18] R. MacDonald, C. Zegura, B. R. Shapiro, J. Borenstein, and E. Zegura, "Developing Community Support for Computing Ethics Teaching Assistants," in *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, in SIGCSE 2023. New York, NY, USA: Association for Computing Machinery, Mar. 2023, pp. 360–366. doi: 10.1145/3545945.3569844.
- [19] C. Zegura, B. R. Shapiro, R. MacDonald, J. Borenstein, and E. Zegura, "'Moment to Moment': A Situated View of Teaching Ethics from the Perspective of Computing Ethics Teaching Assistants," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg Germany: ACM, Apr. 2023, pp. 1–15. doi: 10.1145/3544548.3581572.
- [20] G. Shao, J. P. Quintana, W. Zakharov, S. Purzer, and E. Kim, "Exploring potential roles of academic libraries in undergraduate data science education curriculum development," *J. Acad. Librariansh.*, vol. 47, no. 2, p. 102320, Mar. 2021, doi: 10.1016/j.acalib.2021.102320.
- [21] P. Anderson, J. Bowring, R. McCauley, G. Pothering, and C. Starr, "An undergraduate degree in data science: curriculum and a decade of implementation experience," in *Proceedings of the 45th ACM technical symposium on Computer science education*, in SIGCSE '14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 145–150. doi: 10.1145/2538862.2538936.
- [22] I. Bile Hassan *et al.*, "data science Curriculum Design: A Case Study," in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, in SIGCSE '21. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 529–534. doi: 10.1145/3408877.3432443.
- [23] A. Adhikari, J. DeNero, and M. I. Jordan, "Interleaving Computational and Inferential Thinking: data science for Undergraduates at Berkeley," *Harv. Data Sci. Rev.*, vol. 3, no. 2, Apr. 2021, doi: 10.1162/99608f92.cb0fa8d2.
- [24] K. Patel, A. Brooks-Ramirez, R. Dang, B. Adolfo Ventura Benitez, and L. Yan, "Exploration of Undergraduate Teaching Assistant Identity and Teaching Goals in data

- science Courses,” in *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2*, in SIGCSETS 2025. New York, NY, USA: Association for Computing Machinery, Feb. 2025, pp. 1573–1574. doi: 10.1145/3641555.3705179.
- [25] G. Ladson-Billings, “Toward a Theory of Culturally Relevant Pedagogy,” *Am. Educ. Res. J.*, vol. 32, no. 3, pp. 465–491, Sep. 1995, doi: 10.3102/00028312032003465.
- [26] K. Kishimoto, “Anti-racist pedagogy: from faculty’s self-reflection to organizing within and beyond the classroom,” *Race Ethn. Educ.*, vol. 21, no. 4, pp. 540–554, Jul. 2018, doi: 10.1080/13613324.2016.1248824.
- [27] M. Boenig-Liptsin, A. Tanweer, and A. Edmundson, “data science Ethos Lifecycle: Interplay of Ethical Thinking and data science Practice,” *J. Stat. Data Sci. Educ.*, vol. 30, no. 3, pp. 228–240, Sep. 2022, doi: 10.1080/26939169.2022.2089411.
- [28] K. Gutierrez, B. Rymes, and J. Larson, “Script, Counterscript, and Underlife in the Classroom: James Brown versus Brown v. Board of Education,” *Harv. Educ. Rev.*, vol. 65, no. 3, pp. 445–472, Sep. 1995, doi: 10.17763/haer.65.3.r16146n25h4mh384.
- [29] B. Hooks, *Teaching Critical Thinking: Practical Wisdom*, 1st ed. Florence: Taylor and Francis, 2014.
- [30] F. Marbouti, K. J. Rodgers, H. Jung, A. Moon, and H. A. Diefes-Dux, “Factors That Help and Hinder Teaching Assistants’ Ability to Execute Their Responsibilities,” presented at the 2013 ASEE Annual Conference & Exposition, Jun. 2013, p. 23.588.1-23.588.18. Accessed: Apr. 14, 2025. [Online]. Available: <https://peer.asee.org/factors-that-help-and-hinder-teaching-assistants-ability-to-execute-their-responsibilities>