Characterizing social behavior patterns in teaching assistant interactions with students

Joe Olsen,[1,*] Debbie Andres,[2] Nicolette Maggiore,[3] and Charles Ruggieri[4]
[1]*Graduate School of Education, Rutgers University, 10 Seminary Place, New Brunswick, New Jersey 08901, USA*
[2]*Paramus High School, 99 E Century Road, Paramus, New Jersey 07652, USA*
[3]*Department of Chemistry, Tufts University, 62 Talbot Avenue, Medford, Massachusetts 02155, USA*
[4]*Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, New Jersey 08854, USA*

(Received 13 January 2023; accepted 2 August 2023; published 13 September 2023)

Understanding how relationships between instructors and students develop is important for understanding the undergraduate student experience. We expect the development of positive relationships is related to the social practices (e.g., greetings, using names, sympathizing, or empathizing with students) that instructors use in the course of normal classroom interactions with students. We recorded interactions between instructors and students in remote synchronous online physics problem-solving sessions and surveyed students about their perceptions of their instructors. We selected the highest-rated instructor and lowest-rated instructor in our sample and identified social practices in their conversations with students. We first characterized the frequency of social practice usage by each instructor in their conversations with students. We find that both instructors relied on a set of core social practices in most conversations with students, but that our higher-rated instructor used comparatively more positive commentary and sympathizing or empathizing behaviors than our lower-rated instructor. In comparison, our lower-rated instructor engaged in more negative commentary. Using network analysis, we then explored patterns in co-occurrences of social practices used by each instructor moment-to-moment in conversations and compared the instructors' social practice network patterns. We find that our higher rated-instructor used a greater variety of social practices during moment-to-moment interactions with students, while our lower-rated instructor spent most of his time focused on classroom business. We suggest that professional development for instructors should include guidance on how messages are delivered in classes and encourage the use of high-impact social practices to foster positive relationships with students.

DOI: 10.1103/PhysRevPhysEducRes.19.020129

I. INTRODUCTION

Instructors of undergraduate physics courses represent important figures in the lives of undergraduate students and will extensively shape how students interact with the content of physics and the physics community within their institution. Hence, research on the impacts of teacher-student relationships (TSRs) in physics at the university level is needed, but the literature on TSRs at this level remains underdeveloped. Most work on TSRs is primarily conducted in the K-12 context, where teacher-student relationships are qualitatively different from the undergraduate level [1]. Addressing this gap in the research at the university level is particularly important in science, technology, engineering, and mathematics (STEM), as research suggests that positive relationships with professors are related to improved course grades in STEM [2,3] and retention in STEM majors [3–5]. Careful study of how physics instructors build relationships with students can produce recommendations that may be particularly useful for instructors in the physical sciences.

We posit that the development of TSRs may be related to the social practices employed by instructors in everyday *microinteractions* [6] within physics classrooms. Here, *microinteractions* refer to the small, everyday interactions that occur between two or more people in their usual contexts of interaction. Studying the cumulative effect of these interactions should thus account for the ways that people experience their social connections [6]. We aim to corroborate this perspective with a study of micro-interactions in undergraduate physics classrooms. We sought to characterize the social practices embedded in everyday classroom interactions and examine the relationship between social practices and the strength of TSRs in undergraduate physics classes. We do this with

*joseph.olsen@rutgers.edu

an exploratory study that employs a novel use of network analysis [c.f., [7–9]], to generate visual representations of the contents of everyday classroom interactions between groups of students and individual teaching assistants (TAs) in active learning physics problem-solving sessions.

Our research questions are as follows:

1. What instructor social practices are embedded in conversations with students during active learning physics problem-solving sessions, and how frequently are such practices utilized?
2. In what ways (if at all) can a network analysis visualization of an instructor's social practice co-occurrences enhance our understanding of patterns of social practices during conversations in active learning physics problem-solving sessions?
3. How (if at all) might such social practice frequency and/or co-occurrence patterns contribute to the development of TSRs in active learning physics problem-solving sessions?

This paper presents an analysis of our observations of two different physics instructors, both TAs in an introductory physics course, and attempts to explain why the two instructors differ with respect to TSR quality. We show how social practice usage frequencies and social practice co-occurrence patterns (visualized by network analysis) can reveal differences between the two instructors' in-class practices that may partly explain why the instructors differ in TSR quality. As an exploratory study, this paper raises several questions about the development of TSRs that can be explored in future work and demonstrates how the influence of specific social practices may be tracked using a novel application of a descriptive quantitative method.

II. THEORETICAL FRAMEWORK

A. Relationships and relationship building

As individuals engage with each other over time, their interactions exert a mutual influence on each other. The back-and-forth oscillation of influence is responsible for the formation of *relationships*, which we think of here as a social connection between two or more individuals [10]. From our perspective, all students have a relationship with their instructors simply by merit of the social arrangements prescribed by the university system. These relationships, however, may differ substantially in quality between different student-TA pairs.

By *relationship quality*, we mean to refer to the subjective experience of the connection between individuals. Particular experiences of relationships can range from being very positive to very negative depending on the shared history, present interactions, and anticipated future of the individuals in the relationship [10,11]. The development of positive relationships seems to be a basic need for most people [11]; positive relationships typically contribute to one's sense of well-being, while negative relationships can be quite harmful [1,11,12].

B. Teacher-student relationships

Relationship quality is strongly influenced by the context of the relationships [6,10], warranting individual study of relationships in specific contexts to understand how those relationships function and are developed. Hagenauer and Volet [1] argue that TSRs at the undergraduate level in particular are notably different from their closest counterparts, TSRs at the K-12 level. For example, elementary-aged students exhibit a tendency to depend on their instructors for emotional support in ways that middle and high school students may not [13,14]. At the university level, we might expect an even lower reliance on instructors for emotional support [1].

We follow Hagenauer and Volet in thinking of TSRs at the undergraduate level (hereafter undergraduate TSRs or U-TSRs) as generally consisting of both an *affective* dimension and a *support* dimension. The affective dimension of U-TSRs concerns the emotional attachments formed between instructors and students, whereas the support dimension concerns the efforts made to encourage success at the university (e.g., instructors clearly communicate expectations to students). Instructors who build both affective and supportive bonds with students build strong U-TSRs with students overall.

III. LITERATURE REVIEW

A. Methods for studying TSRs and social interactions

The study of microinteractions has been given some attention in recent research in relationship studies. For example, Blatt and Camden [6] interviewed temporary employees to understand if and how they develop a sense of community at work. The authors found that temporary employees in organizations can develop a strong sense of connection to their coworkers through the accumulation of positive microinteractions with colleagues. For example, one research participant indicated that small greetings and the use of their name in conversation contributed positively to their sense of connection at work.

While we find surveys and interviews to be insightful methods for exploring the question of how relationships develop, we question if these methods can fully capture the potential influence of the microinteractions between instructors and students on U-TSRs during classroom instruction. The psychology literature has documented how the details of specific events are often forgotten while the impression created by these events may durably remain in memory [15,16]. While the development of positive relationships may be related to social practices via microinteractions, we question if the specific details of microinteractions are often forgotten and hence are not always

reported in surveys or interviews despite their impacts being durable.

We argue that classroom observations are a necessary component of exploring relationship development in the undergraduate physics classroom to capture the details of interactions that may not be captured in interviews or on surveys. Studies employing classroom observations that recognize and analyze the social qualities of interactions seem to be almost exclusively conducted in K-12 contexts [17,18]. By contrast, observational studies in classrooms at the undergraduate level commonly focus on the nonsocial qualities of interactions [8,19–22]. For example, Paul and West [23] present the Real-time Instructor Observing Tool (RIOT) as a tool for documenting classroom activities and reflecting on teacher practice. The RIOT documents the class time spent engaging in a variety of categories of interaction. The tool does not attend to the more specific contents of the interactions nor does it document the range of social practices employed in the classroom outside of a select few types of interactions (e.g., chatting with students about nonphysics content). By contrast, we employ a method to more specifically explore the nonphysics content of interactions and characterize the variety in these practices employed by an instructor.

B. Social congruence and teacher immediacy

Other bodies of literature have explored phenomena that may be interpreted from the perspective of TSRs, including the literature on social congruence and on teacher immediacy. We believe these bodies of literature have insights concerning relationship development and briefly review this literature in the subsections below.

1. Social congruence

As a part of a larger research agenda exploring the qualities of facilitators in problem-based learning contexts that impact group functioning, Schmidt and Moust [24] proposed that an instructor's social congruence, or the degree to which the instructor seeks an informal relationship with students and displays an attitude of caring and interest, is an important quality of instructors. Yew and Yong [25] found that instructors who were rated as highly socially congruent were described in a survey using positive adjectives such as kind, caring, and humorous, whereas instructors with low social congruence were described as unapproachable, inflexible, too serious, or too strict. High social congruence was also associated with empathy, care for students, being a good role-model, and being prepared for class. The qualities ascribed to instructors that are highly socially congruent overlap with qualities that are often identified in the K-12 literature as markers of positive relationships [cf., [26]]. Hence, the social congruence literature may provide insights as to what social practices contribute to the development of positive TSRs.

In interviews with medical students and instructors about teaching relationships, Loda *et al.* [27] found specific practices that they claim contribute to a perception of high social congruence. These include having informal discussions during class time (e.g., about cooking or living situations) or sharing personal experiences relevant to course content. Loda *et al.* [28] developed an instrument that sought to identify socially congruent behaviors and found a set of behaviors that were associated with social congruence, including providing helpful and constructive feedback, showing empathy, and taking time for questions. From this work, we can surmise that classroom practices that are characterized by informality, that center on students' thinking, and that acknowledge and make space for student confusion seem likely to impact relationship quality.

2. Teacher immediacy

When engaging with others, certain behaviors can create a sense of immediacy, or perceived physical and psychological closeness, between individuals. Teacher immediacy behaviors have been linked to improved cognitive learning outcomes in students [29]. For the present study, we note that feelings of immediacy are likely to encourage the development of positive relationships, as reducing psychological distance between two people can potentially create fertile ground on which to build a relationship.

Included among the specific immediacy behaviors that have been studied are verbal behaviors such as the use of humor [30], self-disclosure [31], providing feedback to students [32], and body language cues such as smiling or eye contact [33]. Again, we note that some of these behaviors (e.g., humor and self-disclosure) are informal in quality. Other behaviors are centered on recognizing and supporting individuals (e.g., using names and feedback). Since these behaviors are described as creating a sense of psychological closeness, we were sensitive to behaviors with these qualities during our investigation.

IV. PARTICIPANTS AND DATA COLLECTION PROCEDURES

A. Participants and instructional context

The present study focuses on the classroom practices of two TAs at a large, public R1 university on the East Coast of the United States. The instructors, who we identify as H and L, each taught several recitation sections of the same two-semester introductory physics course sequence for nonmajors. At the time of the study, H was a seasoned instructor who had taught in several different active learning settings in a variety of courses over approximately 16 semesters and had observations and teacher training in those course contexts. On the other hand, L was in their fifth semester of teaching in this setting and this course was their only context for teaching experience. Moreover, H had

a vastly greater knowledge of and experience with the common struggles of students as they learn the content. To maintain the anonymity of the instructors, we do not report on the various identities held by both instructors, but we note that H identifies and is perceived as a woman and L identifies and is perceived as a man. The professor for the course requested that all instructors and students keep cameras off at all times during instruction, making the gender difference between the two instructors (perceived through names and voices) the most salient and relevant difference between the two instructors. Furthermore, cameras being off meant that the dominant mode of student-teacher interaction was during class sessions through verbal utterances without any visual context. That is, visual information about the instructor such as facial expressions, body language, and positioning around or among the student groups were not perceived by the students in the moment. In an in-person classroom, however, instructors' facial expressions and body language play a complex role in student learning outcomes and TSR development [34].

The course context for this study addressed topics such as Newtonian mechanics, thermodynamics, and mechanical waves. Instruction at this university relies heavily on lecture-based methods, though most introductory courses in the physics department also include active-learning-based recitations. In a typical recitation, students spend 80 min working in small groups on a selection of physics problems involving concepts introduced in the lecture. TAs are expected to check in on students' progress, facilitate group discussions, and provide assistance when needed.

In the particular course of interest, the course administrator requested that TAs ask all students in their sections at least one content-based question during recitations. Students' individual responses to these questions counted as a part of their participation grade for the recitation. This practice is, to our knowledge, unique to this course's structure at the institution that was the site of data collection. We imagine that this assessment requirement increased the number of conversations between instructors and students as compared to an active problem-solving recitation without this requirement.

At the time of data collection, instruction had moved online due to the COVID-19 pandemic. All instruction was conducted over Zoom. During class time, instructors assigned groups of two to three students to individual breakout rooms where they were instructed to work together on a set of assigned problems. Students were not able to move between their assigned breakout rooms; the expectation was they work with their group members for the entire class period while the instructor roamed between breakout rooms. The shift to online instruction afforded recordings of interactions that were relatively unobtrusive (study team members were not present in the breakout rooms), as all instruction could be recorded through Zoom's built-in recording functions. Though we expect there to be differences in interactions between online and in-person recitations, we suggest that similar methods could be applied to in-person classrooms to understand U-TSR building practices in different contexts. We comment on this further in Sec. VII.

B. Data Collection Procedures

The university's Institutional Review Board (IRB) reviewed and approved all research activities prior to data collection. In the Fall of 2020 and Spring of 2021, we contacted faculty members teaching introductory physics courses at the research site for permission to record classroom interactions and administer a survey in the recitations associated with their courses. After receiving permission from four faculty members overseeing three courses, we then contacted the TAs for each of the recitations and sought their permission to record their classes. Five TAs consented to the recording. For consenting TAs, we recorded a full 80 min recitation in each section the TA was assigned to (typically two to four sections per TA, but sometimes greater than four sections).

At the end of recorded class sessions, we administered to students a validated ten-question survey adapted from Schmidt and Moust [24,25,35]. For each statement on the survey, students rated their agreement with the statements on a five-point Likert scale from strongly agree to strongly disagree. Four of the survey items measured the social congruence of each teaching assistant, which served as a part of our operational measurement of the quality of the TSRs developed in the classroom. These four statements were as follows:

1. "The instructor showed they liked informal contact with us"
2. "The instructor appreciated our efforts"
3. "The instructor showed interest in our personal lives"
4. "I was not afraid to tell the instructor when I did not understand something"

The original survey measured three different constructs. These are social congruence (discussed above); cognitive congruence or the extent to which instructors express themselves in the language of students; and the use of expertise or the extent to which instructors demonstrate expertise in the subject [24]. Large-scale studies have repeatedly found through confirmatory factor analysis that the three expected factors seem to fit the survey well [25,36]. Hence, we feel confident that the subset of survey questions we chose to use for the current study were most closely related to U-TSR development.

We assigned each instructor a social congruence score by first converting Likert scale choices to a numerical score from 0 to 4. For each student who completed the survey and provided consent to participate in the study, we computed a social congruence score by taking a simple sum of the numerical scores for the four items. We recognize that the choice to sum the scores on the four items is questionable

TABLE I. Instructor scores on social congruence survey.

Instructor	Social congruence score (STD Dev)	Number of sections	Number of student responses
T1 (“H”)	12.20 (2.07)	6	51
T2	11.64 (3.81)	4	17
T3	11.64 (3.81)	4	17
T4	10.69 (1.69)	2	26
T5 (“L”)	8.41 (3.31)	2	25

given the ordinal scale of the student responses, but feel this method of aggregating scores is appropriate given that higher scores on any one individual question are meant to indicate higher social congruence [35]. In this way, a sum captures the essential quality that higher scores are more socially congruent. We averaged all student scores for each instructor to assign to each instructor an aggregate social congruence score and chose the highest- and lowest-rated instructor as the subjects of a comparative case study. Again, while the ordinal nature of the Likert scale questions makes averaging across students potentially problematic, we feel the procedure is appropriate here because it allows us to identify instructors that generally scored most highly or lowly on the four items in aggregate, consistent with the purpose of the original instrument. Instructor scores, standard deviations, number of sections surveyed, and total number of student responses across all sections are given in Table I. “H” designates the most socially congruent instructor with a mean score of 12.20 (of a possible score of 16) and standard deviation of 2.07, and “L” designates the least socially congruent instructor with a lower mean of 8.41 but greater spread of student responses indicated by the standard deviation of 3.31.

The two instructors we identified provided instruction in the same course. However, in addition to the sections in the class that H and L taught, H also taught two recitations in a second course covering different content. We decided to only analyze recordings for the classes that the two instructors had in common to control for course lecturer or administrator expectations and effects of content on the instructors’ practices. In total, we recorded and analyzed four class sessions led by H and two class sessions led by L in the course common to both instructors. For scheduling reasons, we were unable to record both instructors in the same week. H and L both covered topics concerning ideal gases but worked on different problem sets with the students in their respective sections. We will comment on the potential effects of the difference in problem sets in Sec. V where appropriate.

To further validate our choice to use the highest- and lowest-rated instructors for our case study, we consulted the end-of-course evaluations for each instructor that are regularly administered by the university. On the end-of-course survey, there was one item that was closely related to the construct of interest for our study, “The instructor [name] had a positive attitude toward assisting all students in understanding course material.” On a five-point Likert scale that was scored from 0 to 4, H had received a mean score of 3.57 ($n = 36$), indicating strong agreement with the above statement, while L had received a mean score of 2.32 ($n = 19$), indicating only very slight agreement with the statement. This result mirrored the results of our own social congruence survey about the two instructors.

When examining the students’ anonymous written comments about the instructors from the end-of-course survey, we found that H had received exclusively positive comments about their performance. H was described as kind, encouraging, helpful, and approachable. One student indicated that “[H] was an amazing instructor. [H] being both personable and knowledgeable allowed for me to be much more open in asking questions.” L, by contrast, had received overwhelmingly negative comments. L was described as someone who “talked down” to students, caused anxiety with his presence, and provided little help. One of L’s students indicated that “I did not want to go to L’s class... I felt L’s attitude was not always positive... I sometimes dread when L comes to ask questions.” From these comments, it is clear that comparing the social practices of H and L in particular can provide insight into the types of social practices that are likely to impact TSRs in both positive and negative ways. Student comments about the two instructors from the end-of-course surveys are provided in Table II.

We are aware that there are systematic biases in instructor ratings along lines of race, gender, and other identity categories [37,38]. Instructor evaluations are related to the students’ aggregate subjective experience with the active learning problem-solving session and their perception of the instructor’s leadership of those sessions, and perceptions and expectations regarding the instructor’s identities will influence those subjective experiences. In our context, all class sessions were taught with cameras off. Therefore, audio-only interactions between teacher and student groups in breakout rooms were the dominant form of communication between students and instructors. Students likely perceived instructor genders on the first day of class through the published name of the instructors and through the instructors’ voices. Any gender biases that may matter in an in-person context may also matter in our virtual context, but the evaluation biases associated with instructors’ visual characteristics are likely to play a diminished role in our setting as compared to in-person class sessions because of the lack of visual information about instructors. Therefore, visual factors are likely not a significant factor in students’ responses in the social congruence survey nor the end-of-semester course survey. The reader should recall that our highest-rated instructor would be perceived as a woman and our lowest-rated instructor would be perceived as a man. If the genders of our instructors were different, we may not have observed

TABLE II. Student responses to the end-of-course survey question, "In what ways, if any, has this course or the instructor [name] encouraged your intellectual growth and progress?".

H comments	L comments
[H] was kind and answered all of my questions	I did not really like the recitation session at all. It was stressful to answer questions each session and it would waste a lot of time of the period and we had barely any time to finish the worksheets
[H] was super helpful and a great TA!	
[H] lead us to the answer not simply giving us it	
[H] was a great TA. [H] was SO approachable when I was confused about a section in physics. [H] did a great job helping me understanding the material.	I felt like [L's] attitude was not always positive and did not answer questions in a way that was straight forward and easy to understand. [L's] corrections on our worksheets however were incredibly helpful.
You have been the best recitation instructor I have ever had. You explain topics really well.	[L] asked us questions that really ensured we understood the material and I found it helpful.
[H] was always approachable and fun to talk to during workshop. [H] always answered my group's questions efficiently. [H's] way of explaining concepts made them easier to understand, and [H] ran workshop very well considering all of the limitations.	Not in anyways really. I felt pressured to ask questions. [L] was nice but I did not want to go to [L's] class.
Always willing to help out in any way either during office hours, during recitation, or in an email	[L] did not encourage intellectual growth and progress. [L] often talked down to the students and asked impossible questions. If the students did not understand a question, [L] did not explain the question in a helpful manner. It seemed like [L] did not want to help students do well during recitation.
[H] was amazing at explaining concepts and answering any questions I had. I always felt comfortable asking questions if I was confused about something because [they] were always willing and eager to help.	I sometimes dread when [L] comes to asks questions if I am not completely sound on the topics. When someone doesn't know something give hints or push them to the answer don't lead them on with what they are thinking or forming their answer if its wrong. Find the good parts even if it is a little piece and build off of it. It sometimes is pointless for me to fumble to answer a question and when I turn down the wrong path away from the correct answer, [L] keeps asking questions that seem like you are on the right path even though it is encouraging the wrong answer
Asks questions to help understand concepts	
Very understanding and encouraging and helped provide questions for us to answer and understand material instead of just giving the answer.	
[H] was a great TA, and [H] was very helpful during recitations with any questions that I had, especially during a rough transition to online courses with COVID–19. However, I would recommend that [H] specifies the amount of points that [H] took off and the reason for participation deductions on each recitation assignment. There were times that I was participating significantly, whether or not [H] was in the Zoom room, and I would receive deductions on my recitation score. Also, I would recommend that [H] cycles around the rooms more often to be able to judge participation more closely. I had to participate extra when [H] was "in the room" for [H] to try to notice that I was participating in that particular instant before [H] leaves 1–2 min later.	No effect was made on my intellectual growth or progress. At times I felt anxiety before class because I was afraid of the questions that [L] would ask later in class. I know that [L] has the best interest for his students but at time when he would explain a topic he would come off as being frustrated by my answer or lack of an answer. Thank you [L] for being patient with me.
[H] was an amazing instructor. [H] being both personable and knowledgeable allowed for me to be much more open in asking questions. Not only that, but [H's] way of explaining was also easy to understand, for the most part.	

the same student responses to our surveys. We comment on this more in Sec. VII.

V. DATA ANALYSIS AND RESULTS

To better organize the paper, we interweave our analytic methods and results. The first subsection describes the development of a coding scheme for social practices. This is followed by the second subsection presenting the coding results. The third subsection describes the use of network analyses to identify patterns in each instructor's use of social practices. Finally, the fourth subsection presents the results of the network analysis.

A. Social practices coding

After identifying our two instructors and the relevant recordings, we transcribed classroom recordings using the transcribing website otter.ai [39] and corrected the transcripts by hand. The research team divided the transcripts into a set of conversations. Each conversation began when an instructor entered a breakout room and initiated their first line of dialogue. A conversation ended when an instructor left a breakout room. If an instructor entered a breakout room but did not speak any lines of dialogue, then we determined that no conversation with an instructor had taken place and any student dialogue was disregarded. Each conversation was reviewed line-by-line to identify social practices in the data. For the present purposes, we identified social practices as verbal utterances within lines of spoken dialogue that were not primarily oriented around communicating content. We chose to only analyze noncontent-oriented social practices for two reasons. First, these practices are most closely related to our measurement of instructors' social congruence. Recall that social congruence measures the extent to which the instructor seeks an informal relationship with students and displays an attitude of caring and interest. Social congruence, as measured by our survey, was found to be meaningfully separable from instructors' use of expertise and cognitive congruence, which measure teacher behaviors around communicating content to students, in prior large-scale studies [25,36]. While there may be relationships between social congruence and other teaching behaviors that fall under the umbrella categories of use of expertise and cognitive congruence, these relationships will need to be explored more in future research and are outside of the scope of this study. Second, our purpose in this study is to show how an analysis of social practices in the moment-to-moment microinteractions with instructors could account for observed differences between instructors in terms of our proxy measurements of U-TSR quality. The literature reviewed in the previous sections supports the notion that noncontent teaching behaviors are related to students' feelings about their instructors (see the prior sections on social congruence and teacher immediacy) and therefore justifies our choice to focus on these practices.

The excerpt below provides an example of an instructor conversational turn that only contains explanations and/or asks questions about content and therefore was identified as not containing any social practices:

Student: We have a question about worksheet two. We need to explain. But we don't know.
H: So when they say kinetic theory of gases, they mean, like, at the level of the individual gas molecules. So this is my container, it has a movable top. So it can move up and down at the level of the individual molecule, how does it hitting the container top and causing the gas to expand relate to the temperature? So what about the motion of the molecule has to do with the temperature?
Student: Temperature moves faster?

Observe that in H's conversational turn, H's contribution exclusively focuses on physics content and therefore does not contain any social practices. On the other hand, instructors' questions or comments about student thoughts and feelings, instructors' classroom management moves, or other noncontent comments were considered social practices, such as in the following exchange:

L: Hey team how are we doing over here?
Student 1: Oh, hello. Good. Yeah.
Student 2: Yeah we're good.
L: What was that, [student 2 name]? Oh, too easy for you guys huh?

In both of L's conversational turns above, L's contributions do not include any physics content. L's first turn consists of a greeting, whereas the second turn contains a question about what a student had said and a question about the group's perceived level of effort.

With Hagenauer and Volet's [1] distinction between affective and supportive dimensions of relationships in mind, we included comments that were significantly related to general aspects of teaching and learning (again, excluding comments primarily containing physics content). Hence, comments that were centered on managing students, explaining aspects of teaching and learning, or discussing other aspects of the university or classroom were still included as social practices for our study. Examples of these types of comments can be found throughout the results and discussion below.

We were sensitive to the identified social practices in the literature described above as we coded our data. However, considering that we were interested in understanding differences in how the two instructors engaged socially with students, we did not begin with an *a priori* coding scheme for interactions. Instead, we inductively

TABLE III. Instructor conversations summary.

Instructor	Number of sections	Number of conversations	Mean conversation time in minutes (range; SD)	Mean number of contributions per conversation (range; SD)	Mean number of unique social practices per conversation (range; SD)
H	4	56	4.87 (0.05–14.2; 3.78)	15.77 (2–50; 11.22)	7.36 (2–13; 2.60)
L	2	25	5.67 (0.33–13.82; 3.64)	22.56 (1–65; 16.19)	7.32 (1–13; 2.62)

generated codes by looking across conversations and identifying both similar and dissimilar social practices between the conversations in our data. Each author of the study independently reviewed the same set of conversations representing approximately 15% of the data, generating codes that described the types of social practices they observed. We aimed to generate a set of codes that described the functions of each social practice within the conversation (e.g., managing students, commenting on thinking, and commenting on teaching behaviors) without inferring instructor intent. After examining a set of interactions, the research team reconvened to discuss differences in coding and generated a refined coding scheme. We refined codes by (i) collapsing codes that were too similar to be distinguished, (ii) eliminating codes that required understanding of instructor's intent, and (iii) generating codes for utterances that could not be captured by an already suggested code. This refined coding scheme was then independently applied by each coder to a new set of conversations and again refined based on the newly coded data. The process was repeated until (i) the coders determined no new codes were being observed, and (ii) the codes were not overlapping in the features they captured. Ultimately, we discovered many of the same social practices that are identified in the literature above, and we will identify and remark on these similarities as a part of our results and discussion.

While coding, we noted that some conversations contained the same type of social practices but differed remarkably in their tone. Hence, social practices were also coded based on the tone of the remark in the context of the larger conversation. When we determined that interactions were likely to be relationship-building experiences between instructors and students, we coded the social practice as positive (+). When we determined that interactions were likely to damage the relationship, we coded the practice as negative (−). Otherwise, interactions were coded as neutral (0) which we think of as maintaining the existing relationship. In following past work that has discussed tone, our judgments about tone were made based on agreement between the authors' individual interpretations of the events as we observed them [e.g., [17]]. When there were disagreements about the tone of a conversation, we resolved disagreements through discussion. We recognize our limited perspective as observers of these interactions and that such judgments about tone may be suspect. To mitigate against over-interpretation, we were conservative as we made judgments about the tone of conversations; we carefully considered the context of statements in our judgments, including student comments before and after an instructor remark, consulted the original audio recordings to listen to the quality of verbal delivery when there was doubt about the interaction, and applied a neutral (0) code to any interactions for which any member of the research team had some doubt or disagreement about the interpretation. We found in many cases that tone was quite easy to agree upon from the text of the conversation, as is exemplified in the following exchange:

Student: Yeah, no, I have to brainstorm that question.
L: Yeah, it's a little tricky, but I think you can get it...

L's conversational turn above provides some encouragement to the student in the form of affirming their ability to arrive at an answer to a tricky question. In cases such as these, we believe that most science educators would agree that the exchange is positive.

B. Social practices results

In total, we observed 56 conversations for H across four observed class sessions and 25 conversations for L across two observed class sessions. Summary statistics describing the two instructors' conversations are presented in Table III. Both instructors spent almost no time in class not engaged with students. Mean conversation time with H and L was 4.87 (median: 4.08) min and 5.67 (median: 5.45) min, respectively. H's mean number of conversational turns per conversation (including all utterances in the conversation) was 15.77 contributions (range: 2–50, $\sigma = 11.22$), whereas L's mean was 22.56 contributions per conversation (range:1–65, $\sigma = 16.19$). Hence, for the conversations in our sample, L's conversations were generally longer and L made more contributions to their conversations.

We identified at least one social practice in all conversations. When collapsing all tone variations into a single type of practice, H's and L's conversations contained an average of 7.36 ($\sigma = 2.60$) and 7.32 ($\sigma = 2.62$) different types of social practices per conversation, respectively. Hence, both instructors included a similar variety of social practices in their conversations. Based on this similarity, the differences between the

TABLE IV. Final codebook of social practices.

Code	Definition	Abbreviation	Subcode	Example excerpts
Student management	TA comments on, inquires about, sets expectations for, or directs the students' working behaviors	M	M+	*H: You're doing just fine. I want to get you all worksheets before it gets too much later. I don't want to be the one holding you up.*
			M0	*H: I wanted to ask [student], real quick about worksheet one. Number one.*
			M−	*L: So, I'm gonna leave you guys because I've been spending a lot of time here…*
Using names	TA uses student names	N	···	*L: Hey [name], how are you doing?*
Greeting	TA provides a formal opening to a conversation	G	···	*H: Hi room 3. How are you?*
Commenting on student thinking	TA comments about students' content-based contributions to the group	T	T+	*H: Perfect. That's a very good answer. Yeah, it was a very tricky question, but you got it.*
			T0	*[Student provides answer] L: Good, yeah*
			T−	*L: Yeah, I think your problem [name] is you're looking for an equation and you're not gonna find one I'll tell you that much.*
Closing	TA provides a formal closing to a conversation	C	···	*H: I'll see you next week.*
Offer of support	TA explicitly offers support to students	OS	···	*H: I have another group asking for help, but you can always do the same.*
Teaching metacommentary	TA comments on or about their own teaching moves or behaviors	TM	TM+	*L: So the reason I'm talking to you about this is because I think a lot of times students don't understand how this thing is even an engine or how it functions, right?*
			TM0	*H: Careful though. So the reason I called it a trap is because be careful what unit you use for T.*
			TM−	*L: I don't care that you're right if you can't tell me why you're right. 'Cause you have a* 50/50 *shot.*
Referencing student contributions	TA references past student contributions to continue the discussion	SC	···	*L: So [name] I want to kind of call it back to what [other name] was talking about, because [other name] was talking about how when you change the volume, that change in volume is the distance in the* $work = fd$ *equation*
Sympathizing/ Empathizing	TA comments indicate understanding of or alignment with student thoughts and/or feelings	SEM	···	*H: It's hard to pick it up really fast. No, I, I totally see what you're saying. I'm glad that [name] was able to come around though.*
Making amends	TA indicates they made some error or mistake, intruded on group members, or otherwise engaged in undesirable behaviors	A	A+	*H: Hi, you were exceedingly patient. How are you doing?*
			A0	*H: It's not my best analogy. I'll say that.*
Acknowledging students	TA acknowledges students' utterances or contributions	K	K+	*Student: We're doing pretty good. L: Glad to hear it.*
			K0	*Student: Thank you so much. H: Yeah, no problem.*

(Table continued)

TABLE IV. *(Continued)*

Code	Definition	Abbreviation	Subcode	Example excerpts
Language	TA uses informal language or colloquialisms	La	···	*H: Oh Damn. [Laughs]. Super Speedy.*
Humor	TA makes lighthearted or joking statement	Hu	···	*L: Because you know it all. [laughter] I gotcha.*
Normalizing	TA compares thinking/working behaviors of students to others not in the group	No	···	*H: You're not the only group to be like, wait, no, we just totally way over thought that.*
Curriculum commentary	TA provides meta commentary on curriculum or on the problem sets for class	CC	···	*H: I feel like they actually could have been a bit clearer on this scenario. They're basically saying you have two tanks.*
Encouragement	TA provides some encouragement to students	E	···	*H: Yeah. So try pushing through that for a little bit. And I think you'll see it.*
Other	Social practices that do not fit in the other categories and do not appear to grouped into other categories	O	···	…

instructors are most likely attributable not to the number of practices but instead to their tone(s) of delivery and patterns in their use of the practices.

The codebook of social practices generated from the data is shown in Table IV, listed in descending order of the total proportion of occurrence across all conversations for both instructors combined. As such, the codes listed first in the table represent the more frequently utilized social practices overall. The code names, definitions, abbreviations, subcodes, and representative excerpts for each code and subcode are given in the table. We describe these codes throughout this Sec. V as we compare the two instructors' social practice use. Code frequencies in each instructor's data appear in Table V. The most frequently occurring practices across both instructors' conversations included greetings (G), closings (C), names (N), commenting on

TABLE V. Code frequencies for H and L.

Code	Subcode	H frequency (number of conversations)	L frequency (number of conversations)
Student management (M)		0.95 (53)	0.96 (25)
	M+	0.27 (15)	0.27 (7)
	M0	0.95 (53)	1.00 (25)
	M−	0 (0)	0.20 (5)
Using names (N)		0.84 (47)	0.96 (24)
Greetings (G)		0.80 (45)	0.84 (21)
Commenting on student thinking (T)		0.79 (44)	0.80 (20)
	T+	0.29 (16)	0.24 (6)
	T0	0.77 (43)	0.80 (20)
	T−	0.00 (0)	0.16 (4)
Closing (C)		0.89 (50)	0.56 (14)
Offer of support (OS)		0.34 (19)	0.48 (12)
Teaching metacommentary (TM)		0.34 (19)	0.48 (12)
	TM+	0.25 (14)	0.04 (1)
	TM0	0.20 (11)	0.44 (11)
	TM−	0 (0)	0.04 (1)
Referencing student contributions (SC)		0.11 (6)	0.24 (6)
	SC+	0.11 (6)	0.20 (5)
	SC−	0 (0)	0.04 (1)

(Table continued)

TABLE V. *(Continued)*

Code	Subcode	H frequency (number of conversations)	L frequency (number of conversations)
Sympathizing or empathizing (SEM)		0.29 (16)	0.04 (1)
Making amends (A)		0.32 (18)	0.60 (15)
	A+	0.04 (2)	0 (0)
	A0	0.30 (17)	0.60 (15)
Acknowledging students (K)		0.76 (42)	0.68 (17)
	K+	0.07 (4)	0.04 (1)
	K0	0.75 (42)	0.68 (17)
Language (La)		0.11 (6)	0.28 (7)
Humor (Hu)		0.16 (9)	0.08 (2)
Normalizing (No)		0.09 (5)	0.04 (1)
Curriculum commentary (CC)		0.25 (14)	0.04 (1)
Encouragement (E)		0.04 (2)	0.08 (2)
Other (O)		0.09 (5)	0 (0)

student thinking (neutral, T0), acknowledging students' comments (neutral, K0), and managing (neutral, M0). These six practices appear to represent core practices for both instructors and will be discussed in the next subsection. The largest differences in the frequencies of social practice used at the conversation level include differences in the use of negative managing (M−), negative forms of commenting on student thinking (T−), closings (C), positive, and neutral forms of teaching metacommentary (TM+, TM0), sympathizing or empathizing (SEM), neutral form of making amends (A0), informal language use (La), and curriculum commentary (CC). We discuss these differences in the second subsection below.

1. Instructor similarities

The frequency analysis shows that there are several social practices employed by both instructors at least once in most conversations. These practices included greetings (G), closings (C), names (N), commenting on student thinking (neutral, T0), acknowledging students' comments (neutral, K0), and managing (neutral, M0). We propose that the two instructors' use of these social practices does not by itself explain the difference in their scores on our social congruence survey or in their student comments. Hence, whatever difference can be observed between these instructors may be attributable to the use of other social practices. Regardless, we describe these practices in more depth below to paint a more holistic picture of the two instructors' teaching.

Generally, the six core practices discussed in this subsection are considered to be good practices in the classroom setting. Instructors who manage their students as they engage in class activities by commenting on, inquiring about, setting expectations for, or directing students' working behaviors (if done without engaging in negative commentary) help students better understand the expectations of the session and whether they are meeting those expectations. Hagenauer and Volet [1] describe these sorts of comments as a part of the *supportive* dimension of TSRs, as opening channels to communicate the keys to success at the university is important for relationship development at the undergraduate level. Additionally, commenting on students' thinking (if done without negative commentary) is a form of real-time feedback that is typically beneficial to students' in-class engagement and learning. As discussed in the literature review above, these types of comments can contribute to a sense of immediacy between instructors and students [32] and are therefore likely to be important for U-TSR development. In the cases of both the management and commenting on thinking practices, we observed that there appears to be no substantial difference in the use of the neutral forms of these two practices between the two instructors. As for the use of names, research suggests that students appreciate when instructors use their names during instruction because they feel respected and recognized within the classroom [40].

Greetings and closings are generally considered polite behavior in American culture, and while we are not aware of research on their impact in the classroom, we believe such practices would only contribute positively to the development of TSRs. It is likely that the greater proportion of greetings (0.84) over closings (0.56) for instructor L is mainly due to the remote synchronous nature of the class session. It is not clear from the students' perspectives if and when an instructor enters the room because students are expected to share a screen while working and there was no audio indicator for the entrance of a new breakout room participant. Instructors must make it explicitly known via an audible utterance when they arrive in a given breakout room (e.g., "Hi everyone, how are we doing?"). On the other hand, sometimes, the conversations in the data ended organically and did not seem to require or benefit from explicit closings. Hence, the closing practice was less present in instructor L's conversations.

Finally, acknowledging student comments demonstrates that instructors are listening to their students and have heard what the students have to say. Generally, these utterances took the form of a short comment, such as a simple "okay" or a "no problem" in response to a student's expression of gratitude. While we are unaware of research done specifically on these types of comments, we imagine that they do not contribute a great deal to positive feelings between instructors and students when the comments are not substantially positive (as is the case with comments labeled K0).

Overall, the expected positive effects from these practices on relationships built between students and their instructors (see the prior section on social congruence and teacher immediacy) might explain why instructor L did not receive lower survey scores than the scores we observed. However, these social practices alone are not enough to make a given instructor well liked and socially aligned with the students, as evidenced by L's scores on our surveys and by L's student comments. We turn to the differences between the two instructors in the next subsection.

2. Instructor differences

While both instructors employed the managing social practice in nearly all conversations, we observed differences between instructors' use of management practices which were positive (M+) or negative (M−). The M+ and M− examples in Table IV provide two contrasting examples. The M+ example provides commentary on students' working progress but the instructor also indicates that they are respectful of the students' time. On the other hand, in the M− case, the instructor is setting expectations that students will continue to work on their own (i.e., without the instructor's help) and then explains they have spent too much time with the students already. The delivery of the message may indicate to students that the instructor no longer wishes to be there, that the students have taken time and attention away from other student groups, or that the students are beyond help within a reasonable time frame (i.e., a lost cause).

To further exemplify the difference we observed, we provide two additional excerpts here. In the M+ example below, H checks on a student who missed part of class due to technical issues:

H: How are you doing, [student]? Did you? I don't know when you might have gotten kicked out. I didn't see it happen. It honestly didn't even tell me that there was somebody back in the waiting room this time.

H's check-in and acknowledgment of the issue communicate concern for the student's learning. We determined the tone of the exchange was positive.

In the M− example below, L has an exchange duration of several turns with a student after asking for the definition of an isobaric process:

Student: ... the isobaric process is when the volume over the temperature is equal... is constant, meaning then that the pressure... since pressure is equal to force over area... it'll also equal like the mass times gravity over the area-

L: [Interrupting student] Hold on one second [student's name], let's not get ahead of ourselves, it's very simple, in an isobaric process... what defines an isobaric process, how about that?

Student: Okay, the type of... like the container that [the gas] is in?

L: No, no, no, because you're giving me a bunch of facts that look, I think they are probably true. I don't really know I'm not sure. But I'm saying like when you have an isobaric process, a process is isobaric because of *this*. What is this?

Student: The pressure is constant, right?

L: Okay, yeah, yeah. The pressure is constant. So the only things you said like I think you said T over V or something, or V over T. I don't know. Maybe it's true. I don't know. But my point is that like isobaric is constant pressure. That's really what I want to get at.

In the example above, the student first attempts to respond to L's content-related question, during which L interrupts and provides feedback to the student that they may be going astray in their response, then L attempts to simplify the original question. The student provides a different response and then instructor L states more bluntly the expectations for the intended response. One should note that the student's contributions thus far are potentially useful for constructing a response, but the instructor does not acknowledge the student's contributions. After the student provides the intended answer, L provides feedback on the student's style of response and not its content, conveying disinterest in exploring their response further because the response did not fit L's initial expectations. All three of L's conversational turns thus contain forms of managing that are negative in tone, including interrupting the student to redirect their thinking in L's first turn and expressing disinterest in the student's contributions in L's second and third turns.

While H's employed managing practices were coded as either M+ or M0, L engaged in managing practices coded as M− in a significant portion of their conversations (about 20%). The observation that H had no instances of M− whereas L did engage in M− practices likely indicates that such relationship-damaging comments on or inquiries about students' working behaviors harm TSRs.

Commenting on thinking is a means by which instructors provide real-time feedback on students' thinking and is valuable for guiding students toward mastery of the learning objectives by providing opportunities to assess and adjust their approach. Both instructors used any form of commenting on thinking social practice with high frequency (H: 0.79, L: 0.80) as shown in Table V, yet we observed differences in the proportion of occurrence of relationship-building and relationship-damaging usage. Below is an example excerpt from instructor L which portrays relationship-damaging commentary on student thinking:

L: Oh, I'm sorry. Um, but from what I heard, I think it was right. I was... I was kind of focusing on something else though.

In this first example, L indicates to a student that they had not listened to the student's response, though they also indicate with some uncertainty that they expect it was correct.

Consider the following line of dialogue as a second example, which occurred after a student indicated they knew a fact was correct because they had recalled it from their lecture:

L: Yeah. Okay. I mean, honestly, [student's name], that's like, my least favorite answer that students give me is that like, when I asked them how you know something, they say, "because [the lecturer] said so."

L indicates to the student that their provided response was not only inadequate but also potentially the worst response the student could have provided from L's perspective. In both examples above and in other similar examples, L's comments are likely to communicate that students' contributions are not worthwhile or are disappointing. Such exchanges are likely to instill in students negative feelings about instructors.

In contrast to the examples above, we present an excerpt from H that shows commentary on thinking which is much more positive in tone. The following line of dialogue occurs after a student provides a response to a content question posed by H:

H: Yes. It's that link between velocity and temperature that I am especially interested in and you cover that so very nice.

In this example, H provides general feedback that the student is correct, specifies aspects of the student's response that were successful, and further comments that the student was highly successful in how they covered the idea. As a whole, the comment provides specific and positive commentary that may generate positive feelings toward the instructor. In Table V, the proportion of occurrence of relationship-building commentary on thinking is similar for both instructors (T+, H: 0.29, L: 0.24), yet, similar to the case for management, the relationship-damaging commentary on thinking differs (T−, H: 0.00, L: 0.16) with instructor H never engaging in a relationship-damaging commentary. Given that both instructors employed T+ social practices in similar proportions and yet only L had instances of the T− practice, this observation further supports the notion that such relationship-damaging utterances may harm TSRs.

Teaching metacommentary (TM), which occurs when an instructor comments on their own teaching behaviors, in general, occurred more frequently in L's than in H's conversations (H: 0.34, L: 0.48) as shown in Table V. We are unaware of research discussing teacher metacommentary of this type in physics classrooms. Teaching metacommentary provides students with insights into the teaching decisions the instructor makes in real time and why the instructor made those decisions. We suggest that such elucidation of thought process and providing rationale for one's teaching moves may help students feel more engaged with the instructor's mindset and reasoning behind instruction choices (that is, having a better understanding of their instructor's behaviors) and are also opportunities for the students to learn more about what motivates the instructor as they teach. Such insights may generally allow for deeper relationship building, especially if those teaching metacommentary utterances are ones that are delivered in positive ways. Below, we provide a sequence of turns in a single conversation illustrating L's relationship-damaging teaching metacommentary:

L: What does that tell you, the work value, the work is done on or by the gas?
Student: By the gas?
L: Okay, now why?
Student: Because... is it because it's smaller than the value of heat?
L: It's not necessarily smaller than the value of heat. That's not necessarily true.
Student: Wait, was I right in saying it's done by the gas? Is that part right?
L: Yes, but... I don't care that you're right if you can't tell me why you're right. 'Cause [sic] you have a 50/50 shot.

In L's example, the student and L go back and forth several times as L attempts to simplify the question. The teaching metacommentary provided in L's last utterance in this excerpt indicates to the student that it does not matter to L if the student is correct or not, because a "yes" or "no" response might only be indicative of an uninformed guess by the student in some cases. L tries to motivate that the reasoning behind the response is more important than the correctness, but the way this is articulated by L and how L

emphasizes that the student's response is misaligned with L's expectations also communicates L's frustration with the student's responses. Furthermore, from the student's perspective, L has not given adequate feedback or guidance on whether the student's original line of thinking was accurate, which the student explicitly prompted L for earlier in the conversation leading up to L's teaching metacommentary. In contrast, below is an excerpt from instructor H's conversations showing positive teaching metacommentary:

H: Did you all have any other questions? I feel like I was a little scarce around your room today. I got caught up in other ones, but... [Trails off]

In the above excerpt, H observes that they spent relatively less time in the students' breakout room than usual and provides commentary as to why this might be the case. H then directs the students to ask any additional questions they might not have had a chance to ask earlier in the session. We interpret the commentary as expressing care for the students' learning and as an attempt to offer an additional feedback opportunity for students, and thus as a positive exchange between students and instructor H.

Instructor L had a greater proportion of neutral teaching metacommentary utterances compared to H (TM0, H: 0.20, L: 0.44). However, in a more extreme case than management and commenting on thinking, not only did instructor H never utter relationship-damaging teaching metacommentary while L did (TM−, H: 0.00, L: 0.04), instructor H also uttered significantly higher proportion of relationship-building teaching metacommentary than instructor L (TM+, H: 0.25, L: 0.04), as shown in Table V. However, given that TM− was used in only one conversation, any impact on TSR development differences may be due to the use of TM+ by H rather than the use of TM− by L.

Offering support (OS), where the instructor explicitly tells students that they are available should the students need any help or have any questions, was generally used less frequently overall compared to other codes. We note, however, that offering support may also contribute to the building of the supportive dimension of U-TSRs, as these comments serve to invite communication between instructors and students. OS appeared in just over one-third of H's conversations and just under half of L's conversations (OS; H: 0.34, L: 0.48). Both H and L frequently offered support to students, yet L engaged in offering support more frequently than H, so if there is any impact of this social practice on the TSR, the impact was not sufficient to overcome the differences between the instructors in other social practices. It is an open question as to the impact offering support has on the TSR as there is not enough evidence here to claim whether it is a practice that builds, damages, or has no impact on TSR development.

The neutral form of making amends (A0; H: 0.32, L: 0.60) consisted of short and mild acknowledgments that an instructor had made an error. Though research has documented teacher responses to student mistakes in the classroom [41,42], we are not aware of research documenting teacher behaviors when the teachers themselves make mistakes. Commonly, A0 codes occurred when an instructor misspoke and apologized before correcting themselves, as in the following excerpt:

L: So [student], you're right on most of what you said. But you're mixing up heat and energy. Uh, sorry heat and temperature.

The neutral form of the making amends code (A0) occurred about twice as frequently in L's conversations when compared to H's conversations. In fact, A0 was one of L's most frequently used practices, as shown in Table V. However, H used two instances of the making amends practice that we coded as positive (A+; H: 0.04, L: 0) given the content of the utterances. In the A+ code example below, the instructor H enters a breakout room after the students in the breakout room had been waiting for assistance for a long period of time:

H: Hi. You have been exceedingly patient. How are you doing?

In this example, the instructor acknowledges that the students have been waiting by complimenting them on their patience as opposed to apologizing for not arriving sooner. The other example of A+ uttered by instructor H is similar in nature. Early in a conversation with students about a particular physics problem, H acknowledges that they are struggling with the solution, and then, after talking out their struggles, articulates appreciation of students' willingness to listen and engage:

H: Here, what are they trying to get at here? This is not the direction to go. So I have the same RMS speed. Why am I struggling with this? I'm struggling with this. I'm going to admit that right now... [Students and instructor discuss problem]
Student: Yeah, so it's this [equation].
H: Alright. Thanks for hearing out my confusion. I'm gonna move on and see how other groups are doing

In a variety of contexts, shifting from providing apologies to similar expressions of appreciation has been shown to improve self-esteem and satisfaction following a service issue [43]. We believe that a similar phenomenon is likely to occur in the present context, but further research in classroom settings is needed.

Referencing student contributions (SC) describes when the instructor refers back to a prior student's utterance with the goal of advancing the discussion around that student's thoughts, typically while interacting directly with a second student in the room. Research on phenomena such

as teachers noticing and leveraging student thinking has revealed various ways that instructors notice and leverage student thinking in the classroom [44,45]. This practice is generally considered critical to student-centered instruction, though research shows that not all instructors notice or leverage student thinking in the classroom, even in very student-centric environments [44]. The SC code was a much more common practice for instructor L compared to H (SC; H: 0.11, L: 0.24) but was generally not frequently used overall. Similar to the case for offering support, even though L used the practice much more than H, this social practice might not explain the observed differences in survey scores between the two instructors. An additional example of referencing student contributions communicated by instructor L is given below:

L: I think [student's name] kind of has this from what I talked to her before. I think [student's name] has a good idea for how to do (b).

In this example, instructor L explicitly references the student's name and provides an acknowledgment that they believe the student comprehends the associated content. L more specifically indicates their belief that the student's prior thoughts will be helpful for the upcoming question, which also assigns greater value to the student's contribution.

Sympathizing or empathizing (SEM) practices describe when the instructor indicates understanding of or alignment with students' thoughts and/or feelings. Sympathizing and empathizing are often overlooked in research in undergraduate STEM education, and the research that does exist does not study this practice on the microinteraction scale in the classroom [c.f., [46]]. SEM was used more frequently by H in about a quarter of conversations, yet only in one instance by L (SEM; H: 0.29, L: 0.04). The function of SEM is to communicate that students' thoughts and feelings are reasonable and perhaps shared by the instructor. An example of H's use of SEM is given below:

Student: Like, some people like don't have time just to do [the homework before class].
H: Yeah, no, I totally get that. Well, I'm a week behind on my grading. Right. So I definitely get it.

In this example, the student articulates the concern that they and other students do not have time to start the homework assignment prior to attending class. For context, the homework always opens at least several days prior to the class meeting day and, while students are encouraged to start the homework before class time, the homework remains open for several days beyond the class meeting, and in practice, many students do not start the homework until after class. Instructor H acknowledges the student's concern, indicates an understanding of the concern, and further communicates alignment with the student's thoughts or feelings by providing an analogous example related to H's time constraints resulting in delays in grading.

The remaining codes corresponding to curriculum commentary (CC), humor (Hu), encouragement (E), normalizing (No), and informal language (La) were either caused by specific curriculum materials issue in H's classes (CC) or were used comparatively rarely (Hu, E, No, La). We present results concerning these codes below, but we feel these are unlikely to explain the difference in observed scores between the two instructors due to their low frequency of occurrence for both instructors.

Curriculum commentary (CC) occurred when instructors provided noncontent-oriented commentary on the class structure or on the physics problems assigned in class. We observed more curriculum commentary in H's conversations than in L's conversations (CC; H: 0.25, L: 0.04). Most curriculum commentary we observed in H's conversations concerned an issue with the wording of a particular problem assigned to students. H repeatedly confirmed with different student groups that the wording of the problem did not make sense, and they instructed students not to complete the problem. The following provides an example:

Student: Worksheet 2. We're confused on number one.
H: Yes. As you should be. This is not phrased well. What did you come up with? Or what... what are you concerned about?

The physics problems used in L's class did not have a similar issue, which likely explains why the CC code was comparatively infrequent in L's conversations (only one instance was observed). We are not sure whether the observed frequency of the CC code is characteristic of H's typical social style in the classroom because of the idiosyncratic issue with the problem sets in some of H's sections. The core problem sets used in the problem-solving sessions led by H and L have been regularly refined for clarity since the 2015–2016 academic year based on in-class teaching experiences. In the year of data collection, the course material designer decided to create multiple versions of each week's materials to mitigate issues with problem sharing across sections and across years. The problem set assigned during H's teaching during the particular week we recorded H's classroom sessions was a newer version of class materials that contained an error. This error was fixed between class sessions and was not present in two out of four of H's classes. We are not sure about the frequency with which these codes would have appeared in a different week of instruction without errors in the assigned problem sets. However, we note that the CC code appeared only twice in the two class sessions using the revised problem sets.

Normalizing (No) describes instances where the instructor compares students' thinking or working behaviors to other students not present in the breakout room, such as compared

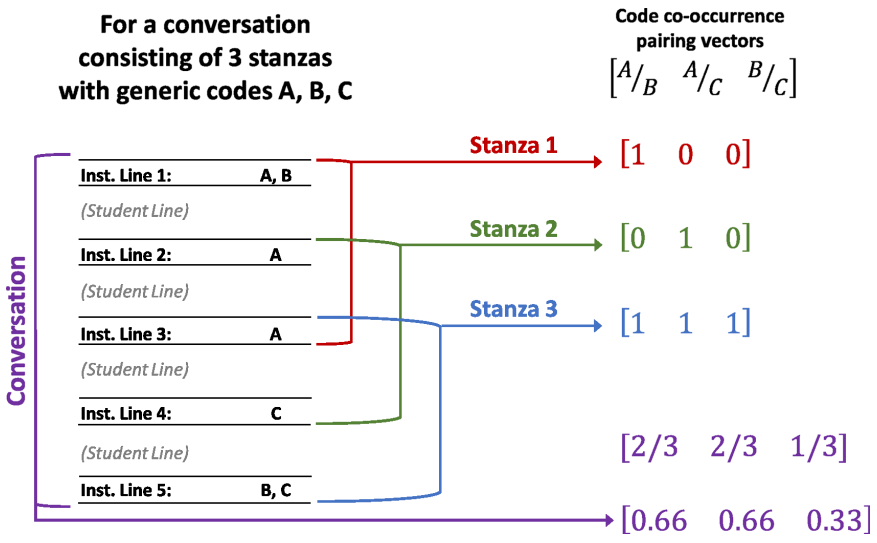

FIG. 1. Schematic of stanza structure and code co-occurrence vector analysis.

to another group in the class. Normalizing was rare across both instructors' conversations (No; H: 0.09, L: 0.04). An example of instructor H's normalizing practice (combined with a curriculum commentary code) is given below:

H: [the worksheet question] should have been more clear on that because you are by far not the only people turned around by this wording.

Normalizing utterances were only used in a positive way, for example, by indicating shared struggles, confusions, questions, or progress that day. Such practices may reduce students' feelings of isolation and imposter syndrome and increase their feelings of community and belonging via commiseration, shared struggles, and so on. Students may experience decreased levels of anxiety or stress after being comforted by the instructor with the knowledge that they were not the only students to encounter a barrier in their understanding or progress.

Use of humor (Hu; H: 0.16, L: 0.08), when done appropriately, is known to improve students' experiences in the classroom and may enhance learning [47]. An example of instructor H's use of humor is given below:

H: Yeah I think your hand might hurt after plugging all of it into the calculator.

Instructors' utterances expressing encouragement (E) served to motivate students to persist and carry on with their efforts and are generally thought of as a productive classroom practice for establishing a positive learning environment. Again, this practice was used rarely (E; H: 0.04, L: 0.08). An example of encouragement utterance by instructor L is given below:

L: Yeah, it's a little tricky, but I think you can get it.

Finally, the use of informal language (La) or colloquialisms occurred more frequently in L's conversations as compared to H's (La; H: 0.11, L: 0.28). Use of informal language or colloquialisms may help socially align students and instructors based on shared informal vocabularies (e.g., "Damn" in the excerpt from Table IV), as such shared vocabularies are often used to signal proximity in a speaker and listener's identities [48]. However, as with prior codes, if there is any positive impact of the use of informal language on TSR development, its impact is not sufficient to overcome L's relationship-damaging practices.

C. Network analysis procedures

Once all conversations were coded, we sought to characterize the moment-by-moment patterns in the micro-interactions between students and instructors. This phase of analysis provided a finer-grained perspective on the interactions between students and the instructors by identifying the most commonly co-occurring social practices within conversations. This analysis provides information about interactions that cannot be captured by frequency analyses alone. While frequency analyses can capture how often certain practices appear, this next phase of the analysis characterized which practices appeared closely together in time and how strong the associations between the practices were for each instructor.

To characterize the moment-by-moment patterns, we adapt a network analysis technique described by Shaffer *et al.* [9], who developed a method known as *epistemic network analysis,* and provided examples of the method as it applies to characterizing the cognitive activities of students engaged in an engineering design problem. The reader should note that we do not follow Shaffer *et al.* exactly but modify their techniques to better suit our purposes. The following paragraphs describe the methods

and computations that we followed to generate our network graphs and a schematic diagram demonstrating our method is presented as Fig. 1. Network graphs were constructed using the igraph package in PYTHON.

To begin this phase of analysis, we further subdivided our conversations into a set of *stanzas*, which consist of a set of three adjacent instructor turns (with student commentary between each adjacent instructor turn; see Fig. 1). These stanzas hence consist of an instructor comment with both the preceding instructor comment and the succeeding instructor comment. Stanzas are inherently nested inside of conversations; stanzas do not span across different conversations, as the dialogue in adjacent conversations is assumed to be unrelated. This assumption is warranted given that every conversation change in the data occurred when instructors moved to a different student group (a different breakout room), limiting the amount of continuity between adjacent conversations. We characterized each stanza by generating a vector describing the co-occurrences of codes within each stanza (Fig. 1, right). In each vector, each coordinate corresponded to a unique pairing of two codes, including tone where applicable (e.g., H/T+ and H/T0 were represented as two different coordinates in the vector). Each vector consisted of a list of 1s and 0s, where 1 indicated a specific pair of codes was represented in the stanza (i.e., both codes were represented on at least one line of the three-line-stanza, though not necessarily on the same line) and 0 indicated that either one (or both) of the codes were not represented in the stanza. By considering stanzas instead of individual lines, we respected the flow of the conversation; immediately adjacent lines in conversations are usually related to each other given the continuous and cooperative nature of a single conversation. Analyzing the data on the single-line level would not capture this feature of conversations and may not reveal patterns in the types of social practices that instructors often use in close proximity to each other. Likewise, analyzing whole conversations consisting of multiple stanzas may mistakenly relate two codes separated by enough conversation turns to be unrelated in their usage.

From these vectors that capture the co-occurrence of pairs of codes in a stanza, we could represent patterns in the social practice contents of conversations through network analysis. Network analysis can characterize the frequency of co-occurrence of specific social practices within each of our instructors' conversations. The social practices identified in the data are represented as nodes in a network graph. Social practices that co-occur within the same conversation are connected by lines. Line thickness represents the frequency of co-occurrence, computed as a simple proportion of the number of stanzas in which two social practices co-occurred to the total number of stanzas in the data (i.e., an average of the coordinates of all of the stanza vectors described above). That is, when two social practices frequently co-occur within stanzas, this frequent co-occurrence corresponds to a greater proportion value and is represented visually with thicker lines connecting the two nodes associated with the relevant social practices. See the top of Fig. 2 for the network diagrams of each instructor. The final diagrams generated by network analysis in this context thus represent patterns in the social practice contents of microinteractions between students and instructors.

To facilitate comparisons between the two instructors, we first generated network graphs from each instructor's data individually and then created a difference graph. Specifically, for each line in the network graphs, we computed the difference in weight between corresponding lines in H's and L's network graphs by subtracting line weights between the two graphs. We then generated a new network graph in which line thickness was determined by these differences, as shown in the bottom left of Fig. 2. Nodes and lines were colored to represent which instructor, H (red) or L (green), had a higher line weight in their network graph. This difference graph provides a visual display of the differences in the two instructors' patterns of social practice use on the microinteraction scale.

D. Network analysis results

Figure 2 presents the co-occurrence network graphs for instructor H (top left, red lines) and instructor L (top right, green lines). The codes are organized by tone (+, −, 0, or no assigned tone). The color schemes serve to easily differentiate between the two instructors' network graphs. The line thickness in the network graphs represent patterns in each instructor's use of co-occurring social practices within stanzas of conversations (as opposed to co-occurrence on the level of conversation). Lines that are thicker indicate that two codes more often appear closely together within an instructor's conversations. A fairly concrete way to interpret the line thickness is as follows: if one were to select a random stanza (i.e., a set of three adjacent instructor conversational turns) from the data, line thickness represents the probability of observing both codes within the stanza. For reference, a line representing a 100% probability of observing both codes within a stanza is presented in the lower right of Fig. 2.

H's network graph consists of a fairly diffuse set of connections between the neutral and positive codes. As expected, the negative codes in the upper right are unconnected to other codes, as H never engaged in the commentary that we considered negative. Note also that normalizing (No) is unconnected to other codes in H's graph. Though H did engage in this practice, she did so infrequently (9% of conversations or five conversations) and she did not use this practice in close proximity to other codes. The fact that the connections in H's network graph appear so diffuse shows that H engages in a large variety of

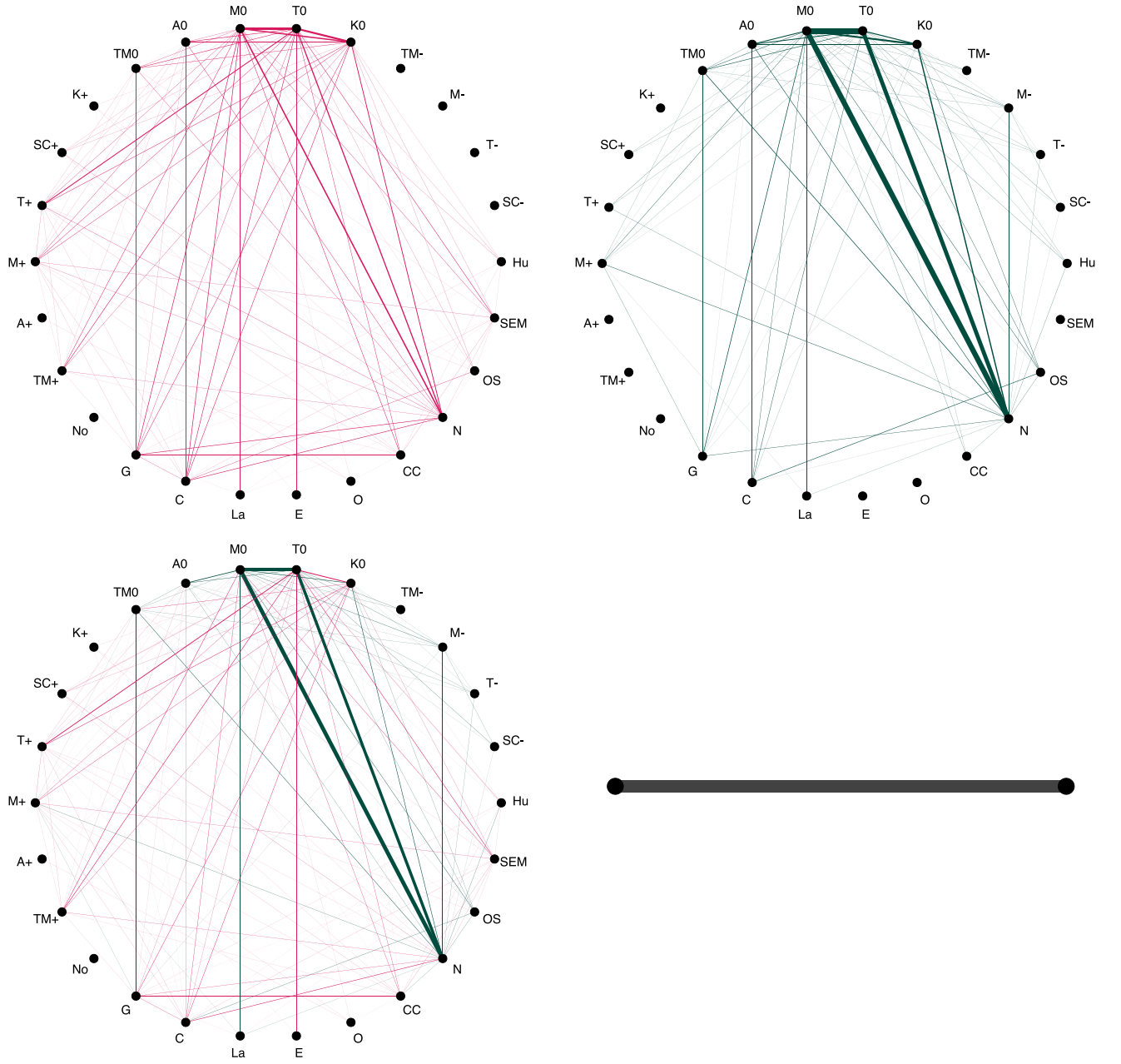

FIG. 2. Network graphs characterizing co-occurrence of social practice use for H (top left, red) and L (top right, green). Line weight represents the frequency with which two practices appeared in the same stanza. If viewing these plots electronically, we recommend zooming in on the plots to better see relative line weights. The reference network graph consisting of two nodes and a single line (bottom right, black) represents two hypothetical codes if they appeared together within every stanza. The bottom left graph was generated by subtracting L's graph from H's graph. In this graph, line weight represents the magnitude of the difference in line weights and the color represents which instructor's graph had a higher weight.

social practices moment-to-moment. The practices that H uses are a mix of both neutral and positive practices. From the graph, we might surmise that one reason H was so well liked among students may stem from the diverse and often positive social practices that characterize H's social interaction style.

L's network graph contains fewer connections than H's graph, suggesting less variety in the moment-to-moment patterns in L's social practices. Note that the negative codes are connected to many different codes (e.g., M− is connected to nine other codes), suggesting that these negative codes were somewhat spread throughout L's conversations versus concentrated in a specific place within a single conversation (which would be visually represented with no connections or very few connections in the graph, see code "No" in H's graph). Like in H's

graph, some nodes are unconnected to the other nodes. These unconnected nodes represent the practices that were not present in L's conversations (with the exception of encouragement [E], which appeared in only one of L's conversations). The most striking visual feature of L's network graph is the strong connections between managing (neutral form, M0), commenting on student thinking (neutral form, T0), and using student names (N). The strong connections between these three codes indicate that these three practices frequently co-occur in L's conversations more than any other combination of codes. The network graph indicates that L's interaction style is much more "down to business" than H's interaction style. Hence, a reason that L may be less generally well liked than H may be attributable to L's more strict focus on completing classroom activities and evaluating student thinking. Indeed, some student comments from L's evaluations concerning the "anxiety-inducing" nature of interacting with L may stem from this orientation toward completing problems and evaluating thinking.

Taking the difference between the two network graphs (Fig. 2, bottom left, red and green lines) highlights the differences described above. In the difference graph, lines that appear in red indicate a higher co-occurrence rate in H's graph versus L's graph, while green lines indicate the opposite. The strong green connections between managing (neutral form, M0), commenting on student thinking (neutral form, T0), and using student names remain in the difference graph and show L's more assessment-focused interaction style versus H's diverse social practice style. Note also that most of the weaker connections in the difference graph appear in red, highlighting H's more diverse use of social practices on a moment-by-moment basis versus L.

VI. DISCUSSION

Above, we described observed differences in H's and L's use of social practices during classroom interactions. We argue that differences in social congruence survey scores and end-of-semester course survey comments may be explained by contrasting H's and L's use of social practices during classroom interactions. We identified a set of social practices that both instructors use commonly in their conversations, including greetings (G), closings (C), names (N), commenting on student thinking (neutral, T0), acknowledging student comments (neutral, K0), and managing (neutral, M0). We expect these practices to have some effects on TSRs, but the evidence we have collected suggests that these practices alone are not sufficient to build strong TSRs. Differences in TSRs in H's and L's classrooms might be explained at least in part by the practices for which there were notable differences between the two instructors. Instructor H never engaged in relationship-damaging forms of any social practices and also more frequently used sympathy or empathy (SEM) practices than instructor L. The combination of the two differences may have resulted in H having developed stronger and more positive TSRs than L. In contrast, instructor L engaged in negative commentary across several social practices (M−, T−, TM−) that were likely damaging to his TSRs. Even though L also engaged in some positive practices and also used certain practices more frequently than H (such as TM0, A0, K0, and OS), those efforts appear to be not enough to overcome the amount of negative commentary in L's conversations (though they may have prevented L from having an even lower score on the surveys).

Moreover, our network analysis technique shows differences between the two instructors in their finer-grained (microinteraction) patterns of social practice use. The salient difference between the two instructors appears to be L's heavy reliance on a combination of managing (neutral form M0), commenting on student thinking (neutral form, T0), and using names (N), while H uses a much more diverse set of practices in their moment-to-moment microinteractions. This difference in H's and L's interaction styles may also in part explain why H's and L's U-TSRs appear to differ in quality.

It is important to recall that instructor H and L have markedly different levels of experience in teaching active learning settings. It is possible that L's relative inexperience in the setting translates to L focusing on what he perceives to be the most critical elements of the session (management, feedback by commenting on thinking, and using names). In contrast, H is more flexible and adaptive in her teaching approach, and she can readily comprehend and address student struggles in real time by drawing upon many years of teaching experience as well as training in and awareness of common student struggles with this content at this level, resulting in conversations that are more relaxed and focused on students' learning as compared to L. In total, it is possible that these differences in teaching experiences influence the instructors' comfort and tendencies in the active learning classroom setting, and therefore influence their choice and tone of the utterances, and ultimately influence the students' perception of these interactions.

VII. LIMITATIONS AND FUTURE WORK

We find that our study supports the notion that an investigation of microinteractions in the classroom can help explain differences in TSR quality between instructors and suggest that repeated and larger-scale study of these interactions might produce even stronger recommendations for practice than those we describe below. However, social practices are extremely challenging to study in rich and complex settings such as active learning classrooms, and there are limits to the claims we can make from the snapshot of classroom practice we have collected for the present study. The specific effects of each of the social

practices in the study are still somewhat of a mystery. It could be the case that some practices, such as the "core" practices we identified above, have positive effects on TSRs, but are somewhat commonplace or expected, and are therefore limited in the impacts they make on students. If this is the case, then there might be "cherry on top" practices (e.g., sympathizing or empathizing) that elevate specific instructors, such as H, far above their peers. It would be informative for future investigations to select some subset of the social practices we identify here and design a targeted intervention to test the effects of specific practices. We also suggest that continued use of network analytic techniques may help scholars identify microinteraction patterns that are related to TSR strength in the classroom.

In our study, we only analyze two instructors and therefore do not have enough data to begin describing these patterns of interaction more broadly. However, our analytic technique did show a marked difference in H's and L's moment-to-moment use of social practices during classroom conversations. A limitation of this study is inherent in our selection of the two instructors at extreme opposite ends of the social congruence survey scores as our case studies. We compare these two instructors because they were the most comparable instructors in our sample; both instructors taught the same class and taught similar content, whereas other instructors in our data taught different classes with different expectations for interactions with students (e.g., one instructor began class with a 10-min minilecture and was not required to initiate conversations with students as H and L were). We leave further analyses of social practices to future studies.

In future work on microinteractions and TSRs, it is worth investigating the grain size at which microinteractions have measurable effects on TSRs. In our study, we characterized microinteraction patterns on the level of conversations and on the level of stanzas (groupings of three adjacent instructor turns). We do not have appropriate data to identify which grain size provides the most explanatory power concerning the TSRs that are developed in the classroom. We note that both levels of analysis (conversation level and stanza level) provide reasonable explanations for the differences in TSR strength between the two instructors. On the conversation level, we noted that differences in the frequencies of specific practices, such as sympathizing and empathizing (SEM) or the use of any of the negative practices, are likely candidates for explaining the differences between H's and L's TSRs. On the other hand, the moment-to-moment patterns that differentiated H and L also provide a reasonable explanation for the observed differences in TSRs. In particular, the network analysis revealed that L's interaction style is much more oriented toward classroom business, which was not immediately apparent from the conversation-level frequency analysis. We find it likely that all of these insights are important for explaining differences in TSR quality but suggest further studies that may provide more confirmatory evidence of this hypothesis.

As mentioned previously, the effects of specific social practices are likely tied to the student perceptions of instructors and of the learning environment, which can be influenced by a variety of factors. These may include instructor professional level (undergraduate instructors, graduate student instructors, faculty, e.g., [49]); the culture of the classroom, university, or cultural background of students [50,51]; racialized and gendered classroom dynamics [52,53]; and many others. It is already well known that there are systematic biases in instructor ratings along lines of race, gender, and other identity categories. Kreitzer and Sweet-Cushman [38] provide a relatively recent review of bias in student ratings of instructors. In our data, we cannot be certain how such dynamics influenced student ratings, and we refrain from providing much demographic information about our instructors for reasons of maintaining anonymity. However, there was an obvious gender difference between the instructors, with H identifying as and being perceived as a woman versus L identifying as and being perceived as a man. Whether H's and L's scores and evaluations would have been closer together or even farther apart if their genders were exchanged between the two is impossible to predict. On the one hand, one might think that H's scores would have been higher if she was not perceived as a woman given that instructor evaluations are often biased against women, yet it is also possible that a man expressing himself in the same ways as H may have been perceived more negatively because of incongruities between the instructor's perceived gender and the instructor's gendered performances [54]. Likewise, L's "down to business" practices enacted by someone perceived as a woman may be more negatively rated by students because of similar incongruities. Further analysis of the relationships between individuals' identities, expression, and use of social practices is needed to understand how these factors together influence relationship development in the classroom.

An additional limitation exists in the purely virtual synchronous nature of our course contexts during the semesters of data collection and the connections to in-person sessions. While the virtual synchronous session audio data were easily recorded, which was ultimately a logistical benefit for our data collection, such sessions are inherently less socially immersive as compared to in-person interactions for a variety of reasons. First, with the instructors and students having cameras off (or at least up to the discretion of the teacher and/or the student), visual information such as body language and facial expression cues, which play a complex role in students' classroom experiences [34], are unavailable to students during conversations in our study. Body language, body positioning,

and facial expressions may enhance (positively or negatively) students' perceptions of the instructor, so the lack of this information in a virtual session means that the TSR development relies solely on audible information. Second, interactions and communication of information between student groups are not possible in the breakout room setup in these courses, so students do not have a chance to listen to or see other groups' progress or struggles nor ask questions of other groups (which is a frequent occurrence in the equivalent in-person setting for these courses). The lack of interaction between student groups creates a greater reliance on the instructor for all feedback and information about the session, which may alter (positively or negatively) the students' perceptions of the instructor during conversations. Third, students in this setting cannot see how the instructor interacts with other groups due to the nature of the isolated breakout rooms which did not allow students to join other groups' breakout rooms. A given student group being unable to witness how the instructor interacts with other groups means that the group has no reference point for the tone and diction of the instructor during conversations relative to other groups. We feel that such differences described above are certainly important when comparing virtual to in-person sessions, yet in both contexts, the instructors' choice of language and tone is ultimately a critical feature of instruction, and the results here certainly apply as a vital piece of in-person teaching strategy. Further analysis and commentary comparing in-person sessions to virtual sessions is beyond the scope of this study.

The data collected in this study consisted of only audible information of the teacher-student interactions. Other modes by which instructors communicate to students or student groups throughout the semester in a typical class setting (virtual or in-person) include email communication, grading practices, and written feedback on graded materials. Based on the authors' several years of experience in this exact class setting for both in-person and virtual sessions, regular email communication between a given student and the instructor is rare, typically only seen for a few students out of the 24 students per section. Emails are more commonly sent to address isolated and sporadic questions or issues, and while one of the most common issues is student absence (e.g., related to illness and emergencies), those issues are accounted for in this class by a separate online form which the instructor does not handle directly. Therefore, we expect email communications to minimally impact a given instructor's aggregate end-of-semester scores or social congruence survey scores, and we do not expect past email communications to play a role in the snapshot of audio interactions gathered and analyzed in this study. Grading practices in this course are standardized by way of a grading rubric used across all sections, for which all instructors received grading consistency training at the start of the semester in addition to spot-checks and feedback opportunities midsemester. We therefore expect the impact of grading practices on the differences of instructors' relationship building to be minimal in this context. Finally, written feedback on graded materials is intended to be seen by all students every week and may directly impact students' perceptions of themselves and their instructor. For example, prior work showed the reportedly positive impact of written assessment feedback on students' self-efficacy in a high school AP physics classroom [55]. While analyzing the breadth, depth, diction, and tone of the instructors' written feedback may provide additional insights into the impact of instructors' written feedback practices on TSR building. The study team does not have access to written feedback data, and we leave such an investigation to future studies.

Finally, we note that the analyses done here were on data obtained in just a few class sessions at a given time during the semester. This limited scope of a few class sessions represents a snapshot of instructors' social practices in time. It would be informative to investigate how these social practices and the associated network analysis graphs change over time and compare them to social congruence ratings over time to help us uncover how these relationships might evolve. We find it likely that the use of specific practices may be time dependent in that some practices may appear more frequently at the early stages of relationship development such as the practices which have a low barrier to use and readily occur in everyday interactions (G, C, N, T, M, SC). Other practices might appear later as a certain level of comfort has been established and have more of an impact on strengthening an existing TSR at later times in the semester (TM+, SEM, E, Hu, No, OS). Studies that explore relationship development over time can show how relationships change with changing patterns in social practices over time.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for open-access publishing costs from NSF DUE/IUSE Award No. 2013315.

[1] G. Hagenauer and S. E. Volet, Teacher–student relationship at university: An important yet under-researched field, Oxf. Ref. Educ. **40**, 370 (2014).

[2] M. Micari and P. Pazos, Connecting to the professor: Impact of the student–faculty relationship in a highly challenging course, Coll. Teach. **60**, 41 (2012).

[3] D. A. Sass, F. Castro-Villarreal, S. Wilkerson, N. Guerra, and J. Sullivan, A structural model for predicting student retention, Rev. High. Educ. **42**, 103 (2018).

[4] B. L. Christe, The importance of faculty-student connections in stem disciplines, J. STEM Educ. **14**, 22 (2013).

[5] C. M. Vogt, Faculty as a critical juncture in student retention and performance in engineering programs, J. Eng. Educ. **97**, 27 (2008).

[6] R. Blatt and C. T. Camden, Positive relationships and cultivating community, in *Exploring Positive Relationships at Work* (Psychology Press, New York, 2017), pp. 243–264.

[7] E. Brewe, L. Kramer, and V. Sawtelle, Investigating student communities with network analysis of interactions in a physics learning center, Phys. Rev. ST Phys. Educ. Res. **8**, 010101 (2012).

[8] K. Commeford, E. Brewe, and A. Traxler, Characterizing active learning environments in physics using network analysis and classroom observations, Phys. Rev. Phys. Educ. Res. **17**, 020136 (2021).

[9] D. W. Shaffer, W. Collier, and A. R. Ruis, A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data, J. Learn. Anal. **3**, 9 (2016).

[10] E. Berscheid, The greening of relationship science, Am. Psychol. **54**, 260 (1999).

[11] B. R. Ragins and J. E. Dutton, Positive relationships at work: An introduction and invitation, in *Exploring Positive Relationships at Work* (Psychology Press, New York, 2017), pp. 2–24.

[12] H. C. Stevenson, Fluttering around the racial tension of trust: Proximal approaches to suspended black student–teacher relationships, Sch. Psychol. Rev. **37**, 354 (2008).

[13] K. R. Wentzel, Social relationships and motivation in middle school: The role of parents, teachers, and peers, J. Educ. Psychol. **90**, 202 (1998).

[14] R. P. Ang, Development and validation of the teacher-student relationship inventory using exploratory and confirmatory factor analysis, J. Exp. Educ. **74**, 55 (2005).

[15] V. F. Reyna and C. J. Brainerd, Fuzzy-trace theory: An interim synthesis, Learn. Individ. Differ. **7**, 1 (1995).

[16] V. F. Reyna, J. C. Corbin, R. B. Weldon, and C. J. Brainerd, How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives, J. Appl. Res. Mem. Cognit. **5**, 1 (2016).

[17] D. Battey, "Good" mathematics teaching for students of color and those in poverty: The importance of relational interactions within instruction, Educ. Stud. Math. **82**, 125 (2013).

[18] J. Allen, A. Gregory, A. Mikami, J. Lun, B. Hamre, and R. Pianta, Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary, Sch. Psychol. Rev. **42**, 76 (2019).

[19] D. G. Wu, A. B. Heim, M. Sundstrom, C. Walsh, and N. Holmes, Instructor interactions in traditional and non-traditional labs, Phys. Rev. Phys. Educ. Res. **18**, 010121 (2022).

[20] M. Sundstrom, D. G. Wu, C. Walsh, A. B. Heim, and N. Holmes, Examining the effects of lab instruction and gender composition on intergroup interaction networks in introductory physics labs, Phys. Rev. Phys. Educ. Res. **18**, 010102 (2022).

[21] E. A. West, C. A. Paul, D. Webb, and W. H. Potter, Variation of instructor-student interactions in an introductory interactive physics course, Phys. Rev. ST Phys. Educ. Res. **9**, 010109 (2013).

[22] V. Dini, H. Sevian, K. Caushi, and R. Orduña Picón, Characterizing the formative assessment enactment of experienced science teachers, Sci. Educ. **104**, 290 (2020).

[23] C. Paul and E. West, Using the real-time instructor observing tool (riot) for reflection on teaching practice, Phys. Teach. **56**, 139 (2018).

[24] H. G. Schmidt and J. H. Moust, What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula, Acad. Med. **70**, 708 (1995).

[25] E. H. Yew and J. J. Yong, Student perceptions of facilitators' social congruence, use of expertise and cognitive congruence in problem-based learning, Instr. Sci. **42**, 795 (2014).

[26] J. H. Stronge, The teacher as a person, in *Qualities of Effective Teachers* (Association for Supervision and Curriculum Development, Alexandria, VA, 2004), Chap. 2, pp. 22–38.

[27] T. Loda, R. Erschens, C. Nikendei, S. Zipfel, and A. Herrmann-Werner, Qualitative analysis of cognitive and social congruence in peer-assisted learning–the perspectives of medical students, student tutors and lecturers, Med. Educ. Online **25**, 1801306 (2020).

[28] T. Loda, R. Erschens, C. Nikendei, K. Giel, F. Junne, S. Zipfel, and A. Herrmann-Werner, A novel instrument of cognitive and social congruence within peer-assisted learning in medical training: Construction of a questionnaire by factor analyses, BMC Med. Educ. **20**, 214 (2020).

[29] D. M. Christophel, The relationships among teacher immediacy behaviors, student motivation, and learning, Commun. Educ. **39**, 323 (1990).

[30] J. Gorham and D. M. Christophel, The relationship of teachers' use of humor in the classroom to immediacy and student learning, Commun. Educ. **39**, 46 (1990).

[31] D. S. Fusani, "Extra-class" communication: Frequency, immediacy, self-disclosure, and satisfaction in student-faculty interaction outside the classroom, J. Appl. Commun. Res. **22**, 232 (1994).

[32] J. Gorham, The relationship between verbal teacher immediacy behaviors and student learning, Commun. Educ. **37**, 40 (1988).

[33] J. A. Sanders and R. L. Wiseman, The effects of verbal and nonverbal teacher immediacy on perceived cognitive, affective, and behavioral learning in the multicultural classroom, Commun. Educ. **39**, 341 (1990).

[34] J. Kerssen-Griep and P. L. Witt, Instructional feedback II: How do instructor immediacy cues and facework tactics

interact to predict student motivation and fairness perceptions?, Commun. Stud. **63**, 498 (2012).
[35] J. H. Moust and H. G. Schmidt, Facilitating small-group learning: A comparison of student and staff tutors' behavior, Instr. Sci. **22**, 287 (1994).
[36] J. C. Williams, W. Alwis, and J. I. Rotgans, Are tutor behaviors in problem-based learning stable? A generalizability study of social congruence, expertise and cognitive congruence, Adv. Health Sci. Educ. **16**, 505 (2011).
[37] G. Potvin and Z. Hazari, Student evaluations of physics teachers: On the stability and persistence of gender bias, Phys. Rev. Phys. Educ. Res. **12**, 020107 (2016).
[38] R. J. Kreitzer and J. Sweet-Cushman, Evaluating student evaluations of teaching: A review of measurement and equity bias in sets and recommendations for ethical reform, J. Acad. Ethics **20**, 73 (2021).
[39] https://otter.ai/.
[40] K. M. Cooper, B. Haney, A. Krieg, and S. E. Brownell, What's in a name? the importance of students perceiving that an instructor knows their names in a high-enrollment biology classroom, CBE Life Sci. Educ. **16**, ar8 (2017).
[41] T. Wan, C. M. Doty, A. A. Geraets, E. K. Saitta, and J. J. Chini, Responding to incorrect ideas: Science graduate teaching assistants' operationalization of error framing and undergraduate students' perception, Int. J. STEM Educ. **10**, 5 (2023).
[42] M. Tulis, Error management behavior in classrooms: Teachers' responses to student mistakes, Teach. Teach. Educ. **33**, 56 (2013).
[43] Y. You, X. Yang, L. Wang, and X. Deng, When and why saying "thank you" is better than saying "sorry" in redressing service failures: The role of self-esteem, J. Market. **84**, 133 (2020).
[44] J. Gehrtz, M. Brantner, and T. C. Andrews, How are undergraduate stem instructors leveraging student thinking?, Int. J. STEM Educ. **9**, 18 (2022).
[45] V. R. Jacobs, L. L. Lamb, and R. A. Philipp, Professional noticing of children's mathematical thinking, J. Res. Math. Educ. **41**, 169 (2010).
[46] V. Arghode, B. Yalvac, and J. Liew, Teacher empathy and science education: A collective case study, Eurasia J. Math. Sci. Technol. Educ. **9**, 89 (2013).
[47] M. B. Wanzer, A. B. Frymier, and J. Irwin, An explanation of the relationship between instructor humor and student learning: Instructional humor processing theory, Commun. Educ. **59**, 1 (2010).
[48] A. Bell, Language style as audience design, Lang. Soc. **13**, 145 (1984).
[49] R. G. Drake, Why should faculty be involved in supplemental instruction?, Coll. Teach. **59**, 135 (2011).
[50] L. V. I. Neal, A. D. McCray, G. Webb-Johnson, and S. T. Bridgest, The effects of african american movement styles on teachers' perceptions and reactions, J. Special Educ. **37**, 49 (2003).
[51] A. M. Villegas and T. Lucas, Preparing culturally responsive teachers: Rethinking the curriculum, J. Teach. Educ. **53**, 20 (2002).
[52] D. Battey and L. A. Leyva, A framework for understanding whiteness in mathematics education, J. Urban Math. Educ. **9**, 49 (2016).
[53] D. Battey, K. Amman, L. A. Leyva, N. Hyland, and E. W. McMichael, Racialized and gendered labor in students' responses to precalculus and calculus instruction, J. Res. Math. Educ. **53**, 94 (2022).
[54] J. Butler, *Gender Trouble: Feminism and the Subversion of Identity* (Routledge, London, 1999).
[55] M. R. Stoeckel and G. H. Roehrig, Gender differences in classroom experiences impacting self-efficacy in an AP Physics 1 classroom, Phys. Rev. Phys. Educ. Res. **17**, 020102 (2021).