ORIGINAL RESEARCH

Fostering noticing of classroom discussion features through analysis of contrasting cases

Kristen P. Blair[1] · Leslie C. Banes[2] · Lee Martin[3]

Abstract
Productive classroom discussion has been shown to support student learning across academic domains. Facilitating successful discussion hinges on the teacher's ability to make adept in-the-moment observations of various aspects of student talk and classroom dynamics. In two studies, we explore a pedagogical intervention using contrasting cases to support novice teachers in learning to notice key features of classroom discussion. Study 1 involves preservice teachers in a bilingual teaching methods course in a university-based credential program. Study 2 involves undergraduates in an education psychology course, many of whom are prospective teachers. Study participants engaged in analyzing transcript-based contrasting cases of discussion vignettes as they collaboratively developed guiding principles for effective class discussion. Data include pre- and post-instruction video noticing task reflections, principles identified, and transcribed partner discussions during the activity. Post-instruction, learners displayed increased student-centered noticing when watching videos of classroom discussions. Additionally, there was increased awareness of the absence of productive features or missed opportunities within the discourse. In this proof-of-concept set of studies, we explore the potential of contrasting cases-based activities to help prepare teachers for the complex task of orchestrating discussion by supporting them in learning to notice.

Across disciplines, educators recognize productive class discussions as instrumental to learning (Gee, 2004; Lemke 1990), and studies over the last two decades have consistently demonstrated positive effects on student achievement (Banes et al., 2020; Mercer et al.,

✉ Kristen P. Blair
  kpilner@stanford.edu

1 Graduate School of Education, Stanford University, Stanford, CA, USA

2 College of Education, California State University, Sacramento, Sacramento, CA, USA

3 School of Education, University of California, Davis, Davis, CA, USA

2004; Murphy et al., 2009). Despite this, productive discussion remains relatively rare in K-12 classrooms in the US, especially in schools characterized by high poverty rates and significant numbers of culturally and linguistically diverse learners (Gallimore et al., 2014; Pianta & Hamre, 2009). By productive discussion, we refer to the collaborative co-construction of knowledge among participants sharing responsibility over content and discourse (Reznitskaya, 2012). The quantity of student talk in a discussion is not enough to create a productive discussion. Classroom research has shown the quality of student talk is equally important (Smith & Stein, 2011; Truxaw & DeFranco, 2008). Orchestrating such high-quality student talk creates a high demand on teachers who must attend to several facets of classroom activity simultaneously (Walshaw & Anthony, 2008). Initiating and sustaining a productive discussion has been described as "among the most challenging activities for an instructor," (Davis, 1993), in part, because classroom talk is unpredictable and contingent upon each talk turn shaping and reshaping discourse.

Within this complexity, one way to begin to understand the learning problem is to consider a simpler analog. At one level, the learning problem might be conceptualized as a matching problem, learning to appropriately pair inputs A, B, C and responses X, Y, Z. Teachers learn to map student speech acts that occur as the discussion unfolds (the inputs) to appropriate responses they might make. For example, if a student puts forward a partially-formed idea, the teacher might probe them to elaborate or provide evidence. The challenge is increased because this occurs not only at single utterance level, but at the level of larger talk patterns unfolding during the discussion. For example, if there is a pattern of "popcorn-ing" of unrelated ideas, a teacher might implement a discussion move to encourage students to listen to and build on one another's ideas. A further challenge is that the space of student inputs is infinite, and talk is not labeled with the kind of input it represents. Teachers need to learn to notice across several dimensions to evaluate features of the discussion and decide which, if any, response is appropriate to enact (Barnhart & van Es, 2015). In other words, they need to recognize A, B, C in context.

According to perceptual learning theorist Gibson (1969), learning to notice involves increased precision in recognizing the underlying structure of a phenomenon, such as a new surfer learning to 'read' the features of incoming waves to decide which to pursue and how to position themselves. With this increased precision comes the ability to differentiate situations that one might initially gloss as similar, such as two waves of the same height that might look the same to a beginner, but which peak in different locations, indicating important distinctions in how they will break to the experienced surfer. In the context of learning to notice student talk during a discussion, this might involve two situations where students are speaking, but the quality of talk patterns differs. Novice teachers might notice broadly that students are talking to each other, while more experienced teachers might have learned to differentiate whether there is equality in who is talking or whether the talk includes genuine uptake of ideas. More precise differentiation should enable more finely tuned responses (e.g., Bransford & Schwartz, 1999). While perceptual learning theories generally focus on increased precision in noticing what is present in the stimuli (e.g., Gibson, 1969), teachers also need to notice what is absent. For example, if a student's idea is not acknowledged and taken up, this absence should be noticed and potentially acted on.

Successful mapping additionally requires teachers to develop a flexible repertoire of possible responses, such as talk moves, to draw upon in facilitating student talk - the X, Y, Zs in the analogy above (Michaels & O'Connor, 2015). This is not as straightforward as it sounds

(e.g., Herbel-Eisenmann et al., 2009), and we posit that it also requires a form of learning to notice, in this case learning to differentiate features of the response moves themselves. For example, educators are taught about revoicing as a talk move they might use. Revoicing has been defined as "the reuttering of another person's speech through repetition, expansion, rephrasing, and reporting" (Herbel-Eisenmann et al., 2009, p. 268). In practice, revoicing can take on many forms, from restating what the student said, to rephrasing in academic language, to including additional information not directly stated by the student. It can also serve many functions, from enabling the speaker to agree with or clarify a restatement, to helping students feel heard, to moving an idea along (Herbel-Eisenmann et al., 2009; O'Connor & Michaels, 2019). When observing model teachers enacting revoicing, or enacting it themselves, it is important for teachers to be able to differentiate revoicing from other kinds of talk moves and to notice the features of *how* it is being enacted in a given situation, which will influence the function it serves and the consequences it might have. This is true of the full space of possible response moves.

While useful for highlighting certain aspects of the learning problem, the simplified analogy of a matching problem does not capture other key aspects of learning to lead discussion. Among them is that discussion occurs in social and cultural contexts and is shaped by factors such as teacher and student values and relational histories at personal and societal levels (e.g., Hufferd-Ackles et al., 2004). For example, with respect to values, teachers often report needing to handle multiple goals that can require navigating tradeoffs, especially in classroom settings, where time is always limited. An example tradeoff is wanting to make space for and validate student interpretations of a literary text, while also wanting to ensure students are exposed to and understand more canonical interpretations (Athanases & Sanchez, 2019). These values may not be surfaced for the teacher, and we posit that a third form of noticing required is for teachers to learn to identify and refine their understanding of their own values and priorities and whether a discussion is unfolding in ways that are consistent with those commitments.

While the full challenge of learning to facilitate discussion is complex, in this paper we tackle three aspects we believe are important for developing teachers, building on Gibson's (1969) theory of learning to notice as developing more differentiated perception of the structure of a situation. First is increased differentiation with respect to the structure and features of student talk, including both what is present and what is missing. Second is increased differentiation with respect to the features and dimensions of teachers' potential responses to that talk, such as when teachers observe other teachers or they themselves apply talk moves. Third is an identification and refinement of one's values and commitments with respect to discussion and more differentiated noticing with respect to alignment with those values.

We present a proof of concept of an instructional approach that uses contracting cases (Bransford & Schwartz, 1999) within a designed activity structure to support noticing. In Study 1, we engaged a small group of pre-service teachers and audio recorded their interactions with the learning materials. In study 2, we engaged a larger group of undergraduates, and included an active control condition who completed a more standard transcript annotation activity for comparison. In both studies, we look at shifts in noticing from before to after the learning intervention and explore how the intervention may have encouraged those shifts.

Prior work on classroom discussion

Features of productive discussion

Discussion has been shown to support student learning across content areas and is often lauded as an evidence-based, high-leverage practice for teacher education (e.g., Forzani, 2014; TeachingWorks, 2023). In math, discussion has been shown to support students' motivation and engagement and help students process mathematical content (Cirillo, 2013). In English language arts and literacy, a meta-analysis revealed that many, but not all, approaches to discussion increased students' comprehension and critical thinking (Murphy et al., 2009). In science, discussion has been shown to help make models of scientific thinking available to students (Duschl & Osborne, 2002) and result in improved reasoning, problem-solving, and understanding (Mercer et al., 2004).

In a productive discussion, teachers and students use one another's ideas as resources to build collective knowledge related to instructional goals. Teachers often have an idealized notion of class discussion, but may be unaware of how specific teacher and student actions can facilitate or hinder it (Adler et al., 2003). Across disciplines, several features of productive discussion have been established in the literature. In contrast to initiate, respond, evaluate (IRE) sequences of classroom talk in which the teacher asks a known-answer question, the student replies, and the teacher evaluates (Mehan, 1982), productive discussion includes authentic open-ended questions, uptake promoting cohesive discourse (Nystrand, 1997), and students taking on leading roles. In high-quality discussions, teachers are not positioned as "knowledge holders," but instead students contribute the majority of ideas (Hufferd-Ackles, Fuson, & Sherin, 2014). Research suggests novice teachers may struggle with eliciting rich student explanations, tending instead to provide the explanations themselves (Banes et al., 2018). They also struggle with positioning students to listen to, build on, and value one another's thinking, not only the teacher's (Hakuta et al., 2013). Teachers and students may use revoicing to repeat, verify, or clarify an idea, facilitating a shared understanding (Herbel-Eisenmann et al., 2009). Ideally students engage in "building on" ideas from one talk turn to another and "building up" of central ideas to develop a complete idea across the discussion (Hakuta et al., 2013). Moreover, in productive discussion, teachers engage diverse student voices (Banes et al., 2018) and probe student reasoning with evidence (Michaels & O'Connor, 2012). Based these ideas, we operationalized five categories of features to focus on as we developed the contrasting cases that form the basis of our instruction: *engaging multiple student voices, eliciting and probing student ideas, revoicing*, *uptake and building of student ideas*, *and bringing evidence into explanations*.

While some facets of class discussion may differ across disciplines, notably, what counts as "evidence" in argumentation (Wolfe, 2011), the features described here were selected as foci for this study of teacher noticing because they have been shown to be effective elements of productive class discussion in studies in math, science, history, and English language arts and can be showcased in short discussion transcripts. Note that these are not the only possible features we could have focused on, but they represent a subset known to be important toward the goal of rich student-centered discussion.

Teacher noticing and professional development

Professional development projects have supported in-service teachers' development of discussion practice, including using talk moves (Michaels & O'Connor, 2015), setting norms, and fostering meta-talk within discussions (Kuhn & Zillmer, 2015). More pedagogical innovations are needed that engage future teachers in thinking deeply about the features of effective class discussion and the teacher and student moves that facilitate it. Tools and approaches that support novice teachers in learning to orchestrate productive discussion early in their careers are crucial, especially when they may not have had opportunities to observe or engage in rich class discussion themselves. Walshaw and Anthony (2008) suggest teachers' beliefs about teaching and learning play an important role in developing complex class discussion practices. Thus, learning experiences that highlight what is possible in class discussion and offer opportunities for reflective analysis may be especially useful. Our approach engages preservice and prospective future teachers in refining their noticing across varied discussions while they reflect on their own values about what makes for a productive discussion.

Some approaches to professional development have focused on teacher noticing. Teacher noticing includes attending to features of classroom interactions, reasoning about what was observed, and deciding how to act (Jacob, Lamb, & Phillip, 2010). Much of the research on teacher noticing has come from the domains of mathematics and science education and has focused on teachers' abilities to attend to and interpret students' mathematical or scientific thinking (e.g., van Es & Sherin, 2021). Research suggests that when noticing in videos of classroom interactions, preservice teachers (PSTs) typically focus on superficial features of classroom practice, such as hand raising, following procedures, and staying on task (Star & Strickland, 2008; Erickson, 2011), and they focus more on teachers' actions than students' thinking and learning (Levin et al., 2009). Research in mathematics and science education has shown that extensive professional development with structured supports, such as reflection and feedback (Davis, 2006), use of contrasting examples (Kisa, 2013), and appropriate framing (Star & Strickland, 2008; Barnhart & Van Es, 2014) can shift what PSTs notice, helping them learn to attend to the features that will better position them to enact rigorous and responsive instruction that builds on student thinking in their domains.

In these studies, our aim is to support future teachers' abilities to notice features of productive classroom discussion that may be applicable across content domains, namely the five feature categories described above. We present these studies as a proof of concept that a relatively short, contrasting-cases based approach can lead to shifts in what participants notice while watching videos of classroom discussion in ways consistent with perceptual learning theories and prior research on productive teacher noticing. We describe the instructional approach and rationale behind it in a way that we hope will be generative for others wanting to develop a similar model.

Instructional approach

History of contrasting cases-based instruction

Our approach to helping future teachers develop noticing skills involves an instructional method derived from perceptual learning theories (Gibson, 1969). These theories initially focused on discriminating visual, auditory, or sensory information, such as detecting differences in the loudness of two tones, and have expanded to examine more 'conceptual' differentiation, such as distinguishing the quality of two writing samples (Lin-Siegler et al., 2015) Contrasting cases are juxtaposed examples that are chosen to highlight distinctive features or relationships (Gibson, 1969; Schwartz & Bransford, 1998; Bransford & Schwartz, 1999). Generally, a majority of features are held constant between examples, so that key differences stand out. For example, pairing wines side-by-side can help people learn to perceive subtle differences they might otherwise gloss. The purpose of contrasting cases is to help people learn to notice important features or dimensions so they can develop a more differentiated understanding of a concept or phenomenon. This can help people learn to distinguish one thing from another, recognize what features or elements are important, and better understand conditions of applicability (Schwartz et al., 2016). Contrasting cases have been shown to support such learning across many contexts including physics (Schwartz et al., 2011; Shemwell et al., 2015), statistics (Schwartz & Martin, 2004; Kapur, 2014), mathematics (Rittle-Johnson & Star, 2007), management science (Roelle & Berthold, 2015), and writing (Lin-Siegler et al., 2015).

A few studies have looked at the use of contrasting cases in the context of teacher education. For example, Schenke and Richland (2017) looked at the relationship between PSTs' mathematical content knowledge and their spontaneous use of contrasting examples of student answers in their instruction. Derry et al. (2007) engaged teachers in contrasting their own solutions to mathematics problems to help them learn to notice opportunities for algebraic thinking in student responses. Two studies have used video-based contrasting cases[1]. Nagarajan & Hmelo-Silver (2006) examined effects of different scaffolding prompts, including compare and contrast and a metacognitive prompt, on what undergraduates in an educational psychology course noticed from contrasting videos of formative assessment interactions between a teacher and student. In Kisa (2013) five biology teachers were tasked with comparing and contrasting across two videos, one showing a small group solving a complex science problem with high levels of student thinking and a second showing a group solving the same task with low levels of student thinking. Kisa found shifts in teacher noticing while engaging with the professional development sessions, with teachers making more linkages between teacher actions and student thinking and more interpretations of student thinking.

---

[1] For the purposes of focusing on classroom discussions, video-based instruction has advantages in terms of the richness and ecological validity of the data sources. However, a potential disadvantage for contrasting cases-based instruction is the degree of variability in these naturally occurring classroom situations. There are many differences between classroom contexts, or even between different discussions within the same classroom, that form challenges to presenting tight contrasts that focus attention on particular features, which are hallmarks of contrasting cases as an instructional technique. In this research, we made the choice to engage students with constructed transcripts of classroom discussions, rather than video-based instruction, although we also think there are instances in which video-based examples have merit.

The ways the contrasting cases are constructed and the activities by which learners engage with them has varied across studies, depending on purpose. A common activity involves asking learners to compare and contrast between the cases with the goal that they will notice key distinctions (e.g.,. Lin- Siegler et al., 2015, Kisa, 2013). For example, Roelle and Berthold (2015) had learners engage with two contrasting examples of company production styles. They compared the effects of having the similarities and differences pre-identified for the learners to having the learners identify the similarities and differences themselves. However, Chin et al. (2016) discuss some limitations of compare-and-contrast task demands, in favor of task demands that encourage synthesis across a set of cases.

An illustrative example of a synthesis task comes from the domain of statistics. Schwartz and Martin (2004) developed instruction where high school students were tasked with inventing a mathematical procedure that would capture the reliability of a baseball pitching machine based on data about where the machine's balls hit relative to a target. To aid them in inventing their procedures, students were provided examples of data from different pitching machines that their procedure should be able to handle, as shown in Fig. 1. These carefully chosen examples - the contrasting cases - were selected to help students notice important features of distributions that a solution would need to take into account. For example, comparing across cases, students might realize that Big Bruiser involved a different number of balls, highlighting the need to consider sample size. Comparing Smyths and Fireball highlights the need to differentiate reliability from accuracy, while comparing Ronco to the other cases alerts students to the concept of outliers. Students invented a range of solutions, which they iteratively refined as peers or the teacher pointed out features of the stimuli that the solution did not account for (e.g., different sample sizes). The instructional goal of the activity was to help students notice important features of the problem space that a solution would need to take into account, such that when students are later introduced to a canonical solution (the mean deviation formula), they better understand the mathematical work it is doing. Through the process of iteratively refining their solutions as they related to the features of the stimuli, increased precision in differentiating the stimuli reinforced differentiation of the solutions, and vice versa. This preparatory noticing activity helped students be more adaptive to new kinds of problems that relied on the same underlying concepts compared to groups of comparison students who learned the mean deviation formula in more traditional ways.

In line with the approach taken by Schwartz and Martin (2004), our intervention presents multiple carefully chosen contrasting cases and a task that requires learners to synthesize across them. The cases are transcripts of classroom talk, representing a space of stimuli (features of student talk) and possible responses (features of teacher moves). Learners in our interventions are asked to synthesize across transcripts to generate three principles that identify what they consider the most important features of productive discussion.

A brief note on terminology: although the participants in our studies are simultaneously students at the university and pre-service or prospective future teachers, in the presentation of our study design and data we will refer to them as "learners" or "participants." This choice allows us to reserve the terms "student" and "teacher" for the discursive roles within the transcripts of classroom talk.

Here are four grids showing the results from four different pitching machines. The X represents the target and the black dots represent where different pitches landed. Your task is to invent a procedure for computing a reliability index for each of the pitching machines. There is no single way to do this, but you have to use the same procedure for each machine, so it is a fair comparison between the machines. Write your procedure and the index value you compute for each pitching machine using the grids below.
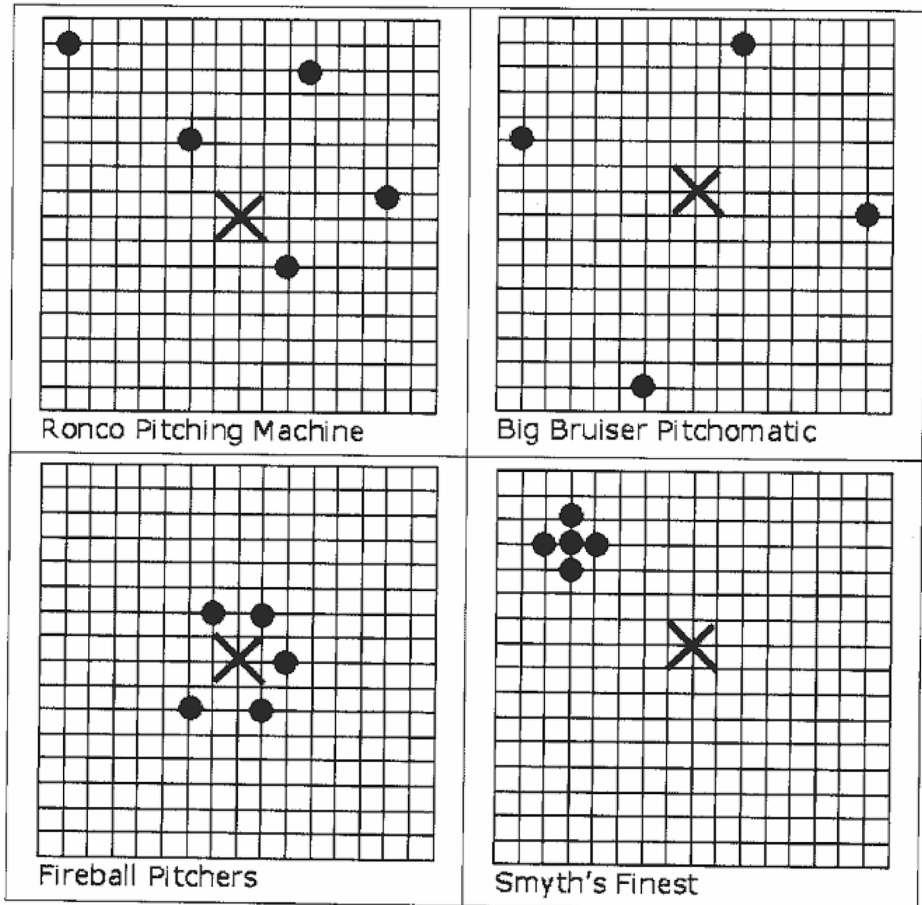

Fig. 1 Contrasting cases used in Schwartz and Martin (2004). Reprinted with permission

Design rationale for the instructional intervention

Design of the activity structure We first describe the design of the overall activity structure and how it relates to the three areas of development that are the focus of the intervention.

a) Increased differentiation with respect to the structure and features of student talk, including both what is present and what is missing.
b) Increased differentiation with respect to the features and dimensions of teachers' potential responses to that talk.
c) Identification and refinement of one's values and commitments with respect to discussion and more differentiated noticing with respect to alignment with those values.

As the basis of the activity, we developed sets of brief, contrasting, fictitious transcripts of classroom interaction. The content of these will be described in more detail below. Their overarching function was to highlight (through contrasts) key features of both student talk and teacher responses to help study participants learn to notice and differentiate these features.

Each study participant received copies of the six transcripts printed on half sheets of paper. They worked with a partner to rank the dialogs from what they thought was the best to the least good example of classroom dialog, however they defined it, and to identify and record two to three principles that guided their ranking. They were instructed that taken together, their principles should explain the ranking of all the dialogs (e.g., for any dialog pair, someone should be able to look at their principles to decide which one they likely thought was better), and they should be able to imagine applying these principles to a new video or transcript of a discussion. There was no single right answer in the ranking or identification of principles. The primary function of the ranking activity was to engage learners in making multiple, close, pairwise comparisons across the cases to help the built-in contrasts stand out. A need to commit to a single ranking among partners ensured that disagreements would be identified, discussed, and resolved. The function of identifying principles was to (a) explicitly name the values related to discussion and (b) discuss how the features of the cases related to those values. Identifying principles helped learners synthesize the ranking of individual cases into a few higher-order principles (Chin et al., 2016). This may support transfer, as learners both differentiate features at the level of the specific stimuli and distill a deep structure that can generalize. Additionally, we hypothesized that identifying what features of a productive discussion should be present would help learners notice instances when those features were absent.

Once they completed the main activity, participants were provided with three additional cases and told that they could modify their principles if they liked. Working with new cases helped learners determine if their principles could handle new examples and refine them if needed.

Finally, participants engaged in a whole class discussion, led by the instructor, about their rankings and their principles. They first entered their group's rankings into a shared document that was projected on screen in front of the class. The instructor then facilitated an open discussion of the rankings and the principles they came up with in their groups, what they meant, and why they were important. This discussion lasted about 10 min. The instructor then gave a 10-minute pre-prepared lecture on features of classroom discourse,

making connections to the just-completed discussion when possible. The rationale for the group discussion and lecture was that activities that ask learners to invent principles or explanations often benefit from an expert recap and create a time-for-telling (Schwartz & Bransford, 1998). Participants may not all notice every relevant feature during the contrasting cases activity. The discussion and expert recap provided an opportunity to see features they missed and consolidate the things they noticed into an explanatory framework that can help organize their understanding. Additionally, the lecture offered a window into research-based, canonical understandings of productive features of discussion, offering an opportunity for learners to refine their values.

Design of the cases The core learning activity made use of a set of six fictitious transcripts of classroom dialog in a middle school science classroom. (See Appendix A for the full set of dialogs cases and activity prompts). Though fictional, the dialogs were developed based on the authors' experiences with teachers and youth in classrooms and thus, represent a range of discussion features that occur in real classrooms. The transcripts presented a classroom discussion focused on whether or not the placement of toxic waste sites is an environmental justice issue, and the six dialogs represented similar but distinct *contrasting cases* of how the dialog could have gone. The dialogs were piloted with a group of teacher educators and revised before they were presented to participants.

As noted above, we developed these transcripts to contrast (and thus make salient) selected discussion features based on a review of the literature, focusing on *engaging multiple student voices, eliciting and probing student ideas, revoicing, uptake and building of student ideas, and bringing evidence into explanations*. For example, across cases, different numbers of students contribute, teachers or students drive the source of ideas, and evidence is brought to bear to varied degrees. Design principles included the need to keep the dialogs short while presenting several variations of the selected discussion features and ensuring dialogs were as realistic as possible. We also wanted to ensure that pairwise comparisons of cases could isolate particular features to make them salient (for example, Dialogs A and F differ only in revoicing), while across the cases, features showed up in different combinations and variations to help participants learn to identify features across instantiations and support discussing and refining principles for ranking (for example, the relative importance of ideas being generated by students and discussions including correct evidence). Appendix B provides a table of the relationship between features and cases.

As an example of the dialog contrasts, consider the first two turns of Dialog A:

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.
T: Okay. Good! Someone on the other side of the room want to share? Can you add something new?
Dialog B had the same opening, but the teacher gave a different response:
S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.
T: So, what I hear you saying is that the location makes it an environmental justice issue because they don't make people sick evenly, right? So, not everyone is impacted

equally by the toxic waste sites. Someone on the other side of the room want to share? Can you add something new?

In this paired example, the teacher in Dialog A affirmed the student's response and moved on, whereas the teacher in Dialog B paraphrased the student's contribution and then offered his or her own language and analysis. This particular contrast was designed to highlight aspects of teacher revoicing.

While we certainly aimed to highlight features meaningful to classroom discussion -- and built the fictitious transcripts from the authors' experiences as teachers and teacher educators to represent kinds of dialog present in real classrooms -- we note that the goal of the present studies was to assess the utility of contrasting cases for promoting noticing. The features contrasted are drawn from the research literature but are not the only features we could have chosen to highlight. For example, we did not in this instance highlight aspects of relational histories among discussion participants or contextual features of the classroom. Among the designed cases, we also did not presuppose which of these features the learners would find most valuable, problematic, and so forth, and this is something we examine in the data. Additionally, the relatively short length of the contrasting cases was chosen to enable the certain features to be made salient without overburdening learners with excessive reading across cases. If this initial proof-of-concept proves fruitful, additional research could consider how choosing different contrasts or longer or shorter cases might influence effectiveness.

Overview of the studies

Study 1 focused on teacher candidates nearing the end of a post baccalaureate teacher credential program to explore their noticing and reasoning around identified features of classroom discussion and examine how the activity might influence their noticing. Study 2 was a replication and extension of Study 1 with a different population (undergraduates -- most of whom were interested in teaching, but who did not necessarily have classroom teaching experience). The population for study 2 was chosen largely because it afforded access to a large enough group of participants to enable a control condition. While Study 1 involved a small sample and no matched control, Study 2 enabled us to introduce an experimental contrast involving a contrasting cases condition and a control condition who completed a transcript annotation task designed to represent a more conventional though still active learning activity.

Based on the research literatures on perceptual learning and teacher noticing, we might expect specific patterns of shifts following instruction. These include:

a) a general shift away from focusing on teacher behaviors toward focusing on student talk and student ideas, consistent with prior documented changes in noticing with increasing expertise (e.g., Levin et al., 2009).
b) increased precision in differentiating aspects of student talk, such as noticing more kinds of features and c) an increase in precision of differentiating the teacher's responses, such as how a talk move is being implemented, which might include noticing more missed opportunities in its implementation.

With respect to noticing of student talk, patterns (a) and (b) would both predict overall increases in student-focused noticing from pre- to post-intervention. With respect to noticing of teacher behaviors and talk moves, pattern a) would predict an overall decrease in teacher-focused noticing as learners shift toward student noticing, while (c) would predict a shift in the quality of teacher-focused noticing, such as noticing more absences.

Study 1 provides an initial test case among advanced pre-service teachers nearing the end of their program. While the population for Study 2 was chosen largely for access, seeing similar effects among a population with less teaching experience would additionally strengthen our proof-of-concept.

Study 1

Participants

Participants included 18 preservice teachers (13 women, 5 men) who were participating in a university course on bilingual teaching methodology and classroom inquiry as part of a post-baccalaureate teacher credential program. Twelve were multiple subject teacher candidates preparing to be elementary school teachers or all content areas, and 6 were single subject candidates preparing to be high school teachers focusing on a specific content area (2 social science, 2 science, 1 English, 1 math). Discussion is relevant in elementary through high school teaching across domains, and all students in the class participated in the learning activity as part of their coursework. All participants were bilingual, taking the course to fulfill requirements to have a bilingual authorization added to their credentials. Seventeen PSTs were Spanish-English bilinguals and one was a Mandarin-English bilingual. All multiple subject PSTs were student teaching in dual language immersion elementary school placements and single subject PSTs were student teaching in English-as-the-language-of instruction high school classrooms. This study took place in the final month of the credential program, and all PSTs had been student teaching for 9 months. All were placed in classrooms with high percentages of culturally and linguistically diverse learners. Their coursework and readings in the credential program had provided some prior exposure to features of effective class discussion, including the importance of eliciting student talk and supporting language learners. Most candidates had not yet engaged with the ideas of uptake or revoicing in class discussion. The quality of class discussion they had observed in their student teaching placements varied.

Materials

Dialog Cases. As previously described, the primary set of cases were six fictitious transcripts of classroom dialog in middle school science that included both teacher and student talk. There was an additional supplemental set of three shorter transcripts that included a few turns of student talk only. We paired these cases with a worksheet that asked participants to rank the cases from what they thought was the worst to best example of productive discussion and to identify three principles that explain their ranking. Full materials can be found in Appendix A. Each dialog case was printed on a piece of 8.5 by 5.5 inch paper, and the worksheet was printed on 8×11 paper. The primary cases ranged from 94 to 232 words

long ($M=160$) and were labeled with the letters A through F. Though the activity was done in pairs, each participant had their own copy of the materials.

Videos for Noticing. As part of an experimental noticing assessment, we asked participants to view and comment on videos of classroom discussion before and after engaging with the dialog cases. Video-based assessments in which teachers detect effective teaching practices are often used as measures of teacher learning (e.g. Kersting, 2008; Santagata & Guarino, 2011; Wiens et al., 2013) and have been shown to be able predict the effectiveness of teachers' pedagogical implementation and student achievement (Kersting et al., 2012). For our video noticing task, we selected four 90 s clips: two clips each from two videos available online through teachingchannel.com and the MA Department of Elementary and Secondary Education[2]. We selected two similar clips from each video, which were counterbalanced, so participants could comment on different clips at pre and post, rather than commenting twice on the same clip in a short span of time, which may have influenced responses. These videos were selected because they were publicly available, had good quality audio and video, and showed students actively participating in a discussion. We note that these conditions were challenging to meet, and we make no claim that these videos were optimal, but selected them only to provide opportunities for noticing. The contrasting cases instructional activity used written transcripts, and so videos of classroom discussion would represent transfer to a more authentic context. The videos differed in the age of the students featured (middle school vs. high school), the ways in which the teachers facilitated discussion (more teacher-driven in the case of the middle school video), and the ways that discussion unfolded (e.g., in patterns of teacher-student vs. student-student talk, with more teacher-student talk in the middle school example). Our intent was to select videos that would allow learners to engage in noticing of the uptake of ideas and other features emphasized in the transcript cases. We refer to the middle school video clips as MS1 and MS2, and the high school clips as HS1 and HS2, with 1 and 2 referring to time order in which they were pulled from the larger video.

Procedure

Preservice teacher participants completed activities as part of the last 90 min of a session of their regular bilingual teaching methods and inquiry course. All participants completed activities as part of regular coursework, but were also given the opportunity to volunteer as research participants, allowing us to collect data on their participation. Everyone in the class consented to be part of the study.

Participants first completed the Video Noticing Task. To begin, we told them, "In a class discussion, when someone says something, there are a variety of ways their idea can be taken up by others. It can be built on, repeated, paraphrased, disagreed with, and so on." We told them their task was to "write down the things you noticed about the ways the teacher and the students take up each other's ideas." Each participant watched a 90 s video clip (MS 1 or 2) and had 3 min to write comments on what they noticed. They then repeated this process for a second clip (HS 1 or 2). To ensure even distribution of each version of the video clips at pre and post, half of the participants were randomly assigned to watch Clips MS1 and HS2, and half were assigned to watch Clip MS2 and HS1. Participants completed this

---

[2] MS videos: https://www.youtube.com/watch?v=Sl72MXNBEZ4&feature=youtu.be&t=205 HS videos: https://learn.teachingchannel.com/video/inquiry-based-teaching-discussing-literature.

activity individually on a laptop while wearing headphones. There was no debriefing on the noticing activity.

Participants then moved to work on the Contrasting Cases activity in pairs. Participants had 30 min to complete this activity. We audio recorded their conversations. They then participated in a short discussion and lecture (approximately 10 min each, described previously) focused on productive classroom discussion.

Finally, participants completed the Video Noticing Task a second time, as a post-test. The process was identical except that participants who previously watched clips MS1 and HS2 now watched MS2 and HS1, and vice versa.

Coding process

We began our analysis with an inductive coding pass of participant responses on the Video Noticing Tasks. We proceeded iteratively through an open coding pass, where all members of the research team completed open codes on a randomly selected subset of the data. We then collaboratively discussed our open codes, debating typical and boundary cases, before reaching consensus on a smaller set of axial codes. Our axial codes included five codes focused on noticing teacher talk and six codes focused on student talk (See Table 1). Based on prior research, we were a priori interested in differences between the teacher and student-focused categories. Additionally, we found that participants sometimes commented both on the presence of a feature (e.g., "the students make points and other students build on those points."), and sometimes actively commented on the absence of a feature or noted a missed opportunity (e.g., "there is a lack of building on a topic from one student to another. Instead, it seems like each new student is offering a new idea to the table"). As such, for each of our eleven codes, we also included an additional tag to indicate if the code was noted as "present" "absent" or both. We call this dimension valence. We did not have a priori expectations about presence vs. absence codes, but wanted to examine whether there were differences in shifts between them, as we hypothesized that noticing absences may be more difficult than noticing presence.

After reaching consensus on this coding scheme and developing a codebook with definitions and examples, two senior researchers coded 25% of the video noticing reflections independently and reached interrater reliability of 83% (exact agreement/total codes applied). Discrepancies were then discussed and the coding scheme further refined the definitions in the codebook until both coders reached agreement.

Two junior researchers and one of the senior researchers then applied the coding scheme to the PST responses. In all coding, researchers were blind to whether a response was from the pre or post video noticing activity. Next, all researchers met to compare coding results and discussed areas with low inter-coder reliability to dispel any differences in interpretations of the coding scheme. This process was repeated until all responses had been coded by all researchers and discussed in meetings (Cohen's kappa average final two rounds of coding: 0.83 presence, 0.82 absence).

Table 1 Codes and Examples for Video Noticing Task. Note: Within descriptions, "T" represents teacher and "S" represents student

| Teacher-focused codes	Description	Example
Acknowledge / validate student ideas	Notices T validating or acknowledging student contributions to the discussion, either verbally or nonverbally.	"teacher either commented on the students' responses or acknowledged them through an affirmation head nod (smile, head nod)" (presence)
Revoice rephrase / paraphrase / summarize	Notices T revoicing or paraphrasing student contributions to discussion.	"the teacher was accepting student responses, but not necessarily paraphrasing or asking for another interpretation" (absence)
Open ended / authentic questions/ probes	Notices the nature of T's questioning as either open-ended, or authentic questions that require an explanation of students' thinking and ideas, vs. closed questions that have one right answer often consisting of one or a few words.	"The questions were not that open ended, with the exception of the last one with "dig deeper, what does it mean?" (presence and absence)
Teacher building on ideas	Notices T building on students' contributions either by taking up a student idea and adding to it	"The teacher is rephrasing some of the students' answers and adding to it mentioning other types of emotions the kids in the story may be feeling: jealousy, fear" (presence)
Scaffold peer to peer building	Notices T facilitating the conversation: e.g., asking questions to prompt students to take up or reflect on others' ideas, agree or disagree, provide additional reasoning and/or evidence, or make connections.	"The teacher acknowledged students answers, and followed up by asking a question that built on the previous answer given by the students." (presence)
Student-focused Codes	**Description**	**Example**
Peer to peer response	Notices S-S patterns of interaction in which students respond to one another in contrast to T-S-T-S pattern (all ideas filtered through the teacher).	"Students had an opportunity to respond to each other" (presence)
Student building on ideas	Notices students building on one another's contributions either by adding on or responding to T questions that prompt students to take up or reflect on an idea, agree or disagree, provide additional reasoning and/or evidence, or make connections.	"Students in this clip are thinking about what was previously said, which is a natural flow for the conversation. They reflect on what another peer said, and then were able to respond to that. i.e. "I agree with…" was heard at the end of the clip." (presence)
Student led / student centered	Notices whether the teacher or students are guiding the discussion and who is the primary source of ideas.	"The teacher is guiding the discussion. The students were talking to the teacher, not to each other. Students were answering the teacher's questions" (absence)
Multiple / many student voices involved	Notices whether many/most voices are represented in the discussion vs. discussion dominated by a few students.	"Teacher chooses students from around the room, and doesn't just stick to a few." (presence)
More student talk than teacher talk	Notices the quantity of student talk vs. teacher talk across the discussion.	"She is not spending a lot of time speaking, but rather allowing the flow to go through the student conversations." (presence)
Use evidence / connect to text	Notices students using evidence to support their ideas (either prompted or unprompted by teacher) and/or connecting ideas back to the text.	"The students give examples from the book, so they have claims that are supported by evidence. This makes the discussion strong." (presence)

Results

Pre- to post- shifts on video noticing task

After coding the Video Noticing Task (see Table 1), we examined differences in participants' responses before and after the Contrasting Cases Activity. We first summed the total number of teacher-focused codes in which the presence of a feature was noted (valence=present) and summed the number of teacher-focused codes in which the absence of a feature/missed opportunity was actively noted (valence=absent). We did the same for the student-focused codes. We then conducted a repeated measures ANCOVA with the within-subject factors of time (pre-post), subject-focus (teacher-focused or student-focused), and valence (noting presence or absence of features). Video clip order was included as a covariate. There was no significant main effect of time ($F(1,15)=0.37$, $p>.5$), but there was a significant interaction of time by subject-focus ($F(1,15)=14.01$, $p<.01$, $\eta_p^2=0.48$). Overall the total number of codes noted did not increase significantly from pre to post, but there were shifts in the focus of codes. Video noticing responses showed different patterns of change for student-focused codes than teacher-focused codes (Fig. 2). Follow-up analysis found a marginal increase in the total number of student-focused codes from pre-to-post ($t(16)=-1.77$, $p=.096$, Cohen's $d=0.43$), consistent with one of our hypothesized patterns of shifts. There were no changes in the total number of teacher-focused codes, $t(16)=0.24$, $p>.5$, Cohen's $d=0.06$. However, small sample size limits interpretation of results. Nominally, there was an interaction within the teacher-focused codes, such that the number of codes noting the presence of a feature in the teacher's behavior decreased while those noting absence of a feature or missed opportunity increased, though this interaction was not significant.

We next examined specific codes along the teacher-focused versus student-focused dimension. The following analyses are exploratory, and do not control for multiple comparisons. Though the overall number of teacher-focused codes did not significantly change, participants were marginally more likely to talk about the presence and/or absence of teacher's open-ended questions at post (exact McNemar's test, $p=.06$). Before the contrasting cases
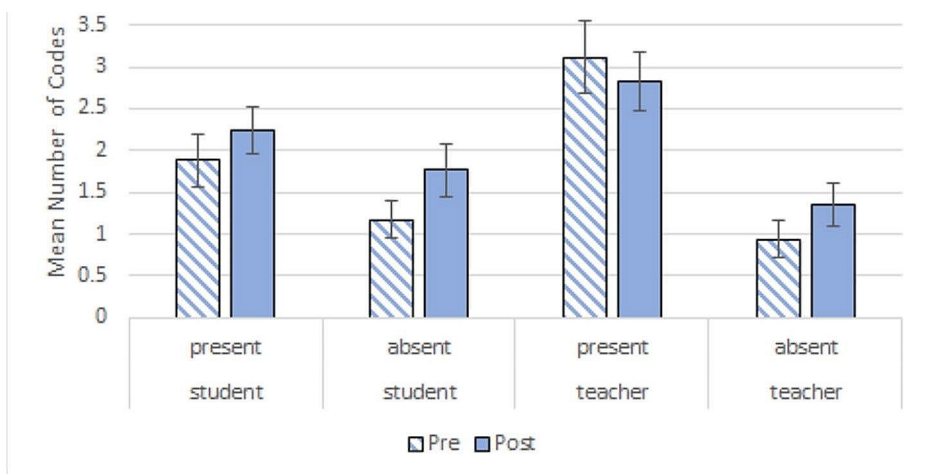
Fig. 2 Mean number of codes pre- and post- Contrasting Cases Activity by subject-focus and valence (noting features as present vs. absent) in Study 1

activities, 65% (11) of PSTs mentioned something coded in this category. After the activities, 95% (16) did. With respect to student-focused codes, there was a nominal trend toward an increase in talking about student-centeredness of the discussion (exact McNemar's test, $p=.12$), with 47% (8) of PSTs mentioning something coded in this category at pre and 75% (13) at post. Participants were more likely to notice the absence of students building on each other's ideas (exact McNemar's test, $p=.03$), mentioned by 6% (1) of PSTs at pre and 41% (7) of PSTs at post.

What happened during the intervention

As pairs ranked cases and developed their principles, they identified and grappled with what they prioritized in a discussion and how it played out in the dialogs. There was variability in how the groups ranked different dialogs, as can be seen in Fig. 3. Participant pairs uniformly rated Dialog E, which was completely student-led, as the best, followed by dialog C, which involved extensive teacher eliciting of student ideas. Ratings of the remaining dialogs showed more variability. Dialog B, referenced earlier, where the teacher did a considerable amount of elaboration while revoicing student comments, was particularly controversial (mean rank 4.11, SD=1.45). Three of the nine groups ranked it as worst (sixth), while five groups ranked it as third best. While groups varied in how they ranked it, all noticed a key feature it was intended to highlight: whether the teacher interjected additional ideas into the
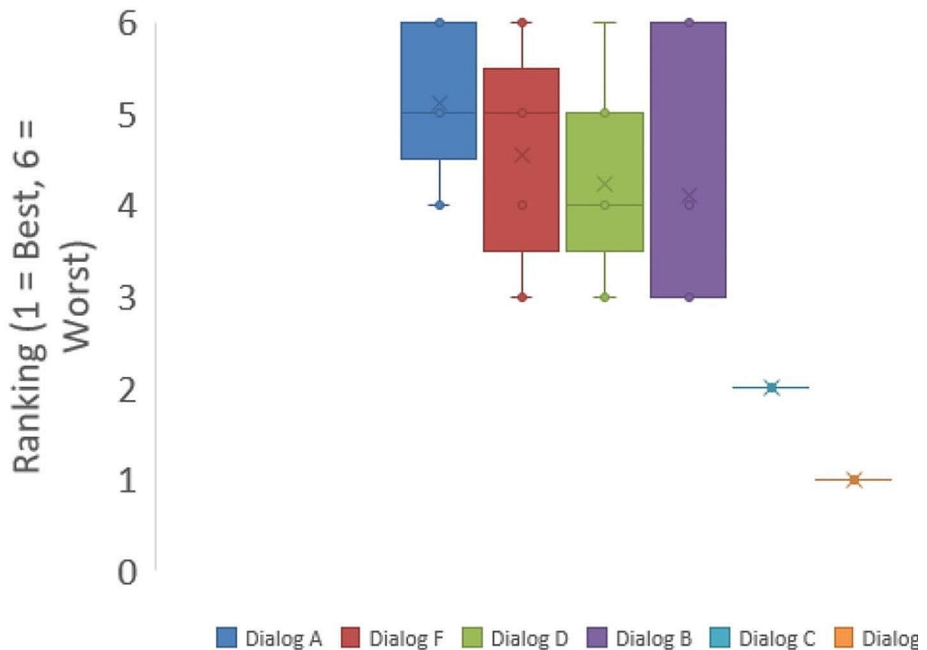

Fig. 3 Ranking of dialogs by preservice teachers in Study 1. A rank of 1 is best, and the dialogs are listed in the order of least good to best (left to right)

conversation when revoicing. The differences in the rankings of Dialog B reflect differences in valuing, rather than differences in noticing the key feature.

To illustrate the process behind the rankings, consider the discussion that one pair of participants, Carla and Maggy, both multiple subject PSTs, had around Dialog B as they discussed teacher rephrasing and revoicing, which they noted can be positive, but can also represent the teacher leading the conversation at the expense of student voice. We selected this audio-transcript as a representative example of how, through ranking and closely comparing the contrasting cases, PSTs reflected on tensions they experienced in their moment-to-moment decision making and their developing beliefs about the goals of productive discussion.

The pair first compared Dialog B to Dialog D, which included many teacher-led short answer questions.

Carla: The teacher also rephrased it in B. I like rephrasing because sometimes you understand things [inaudible]. Yeah, so I like when the teacher rephrases something, because one, it'll probably be louder, and two, it'll be concise, whereas the students tend to ramble sometimes.

Maggy: Yeah.

Carla: So, to me, that was a good thing, when a teacher rephrases. So I like that D and B rephrase. But again, D has no sentences from the students, although the teacher is asking them questions.

Maggy: Yeah, but this is a….

Carla: But it's not high level? What do you think?

Maggy: I just don't like how the teacher is leading them to the… It's kind of a yes and no question. "Do you think that this is bad?" "Yes." "Why?" You know? It's very low level.

Carla: Do you agree that the teacher's doing that in both of these, the teacher's leading the conversation?

Later, this pair contrasted Dialog B with Dialog A, continuing to consider the role of students and the teacher in leading the discussion and contributing ideas, and referring back to their own classroom instruction to further problematize rephrasing.

Carla: For me, it goes back to, do you want students to lead or something? If our goal is for students to lead and cooperate on this question or come up with their own idea, then A is gonna be better than B, but that brings in the fact that A would be better than B [inaudible] because again, the teacher isn't creating the conclusion for them.

Maggy: This is fun! I'm having a whole load of fun, I don't know why.

Carla: Well, 'cause….

Maggy: From this answer the student gave, the teacher is like, "So I hear you're saying that the location makes it an environmental justice issue because they don't make people sick evenly," right? And it's like, of course the student's not gonna say, "No, it's not an environmental justice issue".

Carla: Yeah, right.

Maggy: Well I'd hope they'd say that, but you're right, they probably won't.

Carla: So not everyone is impacted… and then, the teacher's just, you know, making conclusions for the students, so not everyone is affected equally by the toxic waste sites.

Maggy: Yeah.

Carla: So that, "Someone on the other side of the room want to share," they already shared their opinion, or like, supposedly what the student's opinion was, but it sounds like

that's what the teacher thinks, too. And I know when I do that in my classes, when I'm like, "Oh, so do you think that Fidel Castro [was a hero] because he did this and this?" And then the students are like, "Uhhh, yeah, yeah he was a hero."

Maggy: Because of what you already said.

Carla: I don't know, yeah.

Carla: Okay, so then… For me, I feel like A is a little better, even though there's not a lot of talking, they're still coming up with their own conclusions here.

Maggy: Right.

Carla: I don't know. What do you think?

Maggy: Honestly….

Maggy: It really depends on what I want them to take from this discussion, but if we do think that a discussion… 'cause we also have to define how….

Carla: So one of the ways you're defining a discussion is students need to learn how to create their own ideas and put those into a discussion. I think B is worse. A is better than B.

For Carla and Maggy, Dialog B prompted them to consider several important concepts. Comparing Dialog B to Dialog D brought up ideas around rephrasing and high- versus low-level questions. Later, when they compared Dialog B to Dialog A, they talked more about the sources of their ideas and conclusions. After relating the teacher moves in Dialog B to their own experiences leading discussion, they came to the conclusion that, for them, a central component of a discussion was that students are creating their own ideas and coming to their own conclusions, which they then included as a principle for ranking the dialogs. In this example, comparison across the dialogs helped the participants to differentiate their perception of a teacher talk move (revoicing), as well reflect on and refine their own values and commitments when it comes to discussion.

A second exemplar illustrates study participants attending to and differentiating features of student talk across contrasting cases and identifying them in their guiding principles. To illustrate how this may have plausibly influenced noticing, we present a conversation around the three supplemental dialogs (G, H, and I).

Across the three dialogs, student one says the same opening line:

S1: Like White people and people with money don't get sick as much from the toxic waste. So that means it's not fair.

The dialogs differ in student two's response. Dialogs G and H contrast 'I agree because' sentence frame use with responses that differ in degree of uptake and building. Dialog I presents an example of uptake that does not use a sentence frame.

Dialog G.
S2: I agree with David because rich people don't live by the toxic waste sites. The toxic waste sites are mostly near low-income neighborhoods like we can see on the map.
Dialog H:
S2: I agree with David because pollution is really gross, like when there's garbage in the ocean and it kills the dolphins. People shouldn't pollute the earth.
Dialog I:

S2: People shouldn't get punished and get sick more because of where they live. They can't afford to move.

Consider the following dialog between Brandon (single subject English) and Javier (single subject science) as they negotiate meaning and unpack their understanding about what it means for students to "build on each others ideas." While all PSTs negotiated meaning with their partners, this transcript was selected as an exemplar of partners jointly constructing a nuanced understanding of one of their key principles, using examples from the contrasting cases.

Javier: I'm going to add direct responses [to our guiding principles]. Or responses that build on the other.

Brandon: Yeah.

Javier: This one's not a good one because [the student] is not really responding to the other student. He's just like "The ocean's really gross." That's not what [the first student] was saying. He wasn't saying that [pollution] was gross.

Brandon: Well, they're saying that [the placement of toxic waste sites] is not *fair*.

Javier: Yeah.

Brandon: But it's not building off of what he was saying. Yeah. They're just bringing-.

Javier: It's like he wasn't paying attention.

Brandon: -In another point.

...

Brandon: It's like, how well were they able to improve… So it looks like students that were actually able to improve the first person's argument. The first person on the top dialogue was saying "White people and people with money don't get sick as much from toxic waste, so that means it's not fair." And then student two is saying "I agree with David because rich people…".

Javier: Yeah.

Brandon: And then they say that the toxic waste sites are mostly near low income neighborhoods like we can see on the map.

Javier: Yeah.

Brandon: Instead of just saying "that means it's not fair." Yeah.

Javier: Yeah.

Brandon: So that is a big leap, and they are building off of-.

Javier: Yeah, they're directly building off of it. Like the third one-.

Brandon: Yeah, it's a real improvement.

Javier: "People shouldn't get punished." He kind of took it a little further.

….

Javier: Yeah. And the other one's just like "The ocean's gross."

In the above conversation, Brandon and Javier reason that for students to "build on each other's ideas" their conversation turns must be closely related. Brandon and Javier decide, despite the use of marked linking language or sentence frame ("I agree with ___ because…"), Dialog H is actually *not* an example of building on student ideas. Comparing dialogs offered opportunities for Brandon and Javier to deepen their understanding of discussion as they jointly decided responses must "improve" an argument or "take it a little

further" to count as "building." Moreover, they reflect an understanding of how "building" can sometimes serve to fill in gaps ("big leaps") in the class's collective argument by clarifying ideas and adding evidence. They contrast this example of strong student-student uptake to the pseudo-uptake they identified in the first transcript, which may have influenced their attending to student ideas. In this case, as the pair were refining their principles, they were also plausibly refining what they were noticing in the student talk sequence.

In all, the principles that participants generated were related to the contrasts built into the cases. Each of the nine pairs of participants identified three principles, yielding 27 in total. We combined like principles and categorized them into seven groups (guiding principle categories), some of which were teacher-focused, and some of which were student-focused (see Table 2). The most common principle identified by the PST pairs was that the discussion should be student-centered or focused, which is perhaps reflected in the pattern of a larger increase in student-focused noticing codes post-instruction. With a larger sample size in Study 2, we will examine how principles identified during the activity relate to shifts in noticing patterns after instruction.

Study 2

Study 1 showed that the Contrasting Cases Activity led to fruitful conversations among preservice teachers, with evidence of shifts in noticing on the Video Noticing Task. However, a small sample size and lack of a comparison condition limited interpretation of results. In Study 2, we sought to replicate and extend Study 1 with a larger sample and addition of an active control as an experimental contrast. We also moved to a different population of participants, in this case undergraduates who were taking a course on educational psychology. This shift afforded the opportunity to have a sample large enough to include a control condition, and it allowed us to explore how more novice learners, who had not gone through a teacher credential program, would respond to the instructional intervention. Previous surveys of undergraduates in education courses at this university found that more than two-thirds reported teaching as a definite or possible career goal. Because of this, we refer to these undergraduate as prospective teachers, though not all may choose to pursue teaching. Compared to the preservice teachers in Study 1, they had less knowledge and experience about teaching.

Participants

Participants in Study 2 were undergraduates enrolled in an introductory course on educational psychology. Only participants who consented to participate and completed all phases of the study were included in the sample ($N=104$; 86 women, 16 men, 2 non-binary; 5 sophomores, 48 juniors, 51 seniors.) Due to a technical error, for 18 of the participants, the videos in the pre- and post-test video noticing task were not counterbalanced; they saw the same videos at pre and post. Those participants were removed from comparative analysis of pre-post differences, resulting in $N=86$ for these analyses, though their in-class worksheets were still included in process analysis.

Table 2 Most frequent categories of guiding principles in Study 1

Guiding Principle Category	Example of principle	Example dialog contrast related to the principle	Study 1 % PSTs pairs who mentioned (count) $N=9$
Discussion is student led or student centered	"How well was the discussion based on the student's ideas and how well were the students having dialogue with each other."	Dialog C vs. B	78% (7)
Teacher asking open-ended or deeper probing questions	"I also took into consideration if the teacher was using close-ended questions or not. If they were, I rated lesson less good because in my opinion it doesn't give room for students to express themselves or dialogue"	Dialog C vs. D	55% (5)
Evidence or data brought into discussion	"Using evidence (map/graph/articles) to support their claims"	Dialog C vs. F	44% (4)
Student ideas validated (by teacher or other students)	"Validating each other's ideas and using student names. 'Let me try to summarize' 'What I hear you saying is___' 'David and Jose' - use student names to give credit/validate said inputs. 'I'd like to add to David'"	Dialog B vs. A	44% (4)
Students building on each other's ideas	"Students building on each other's ideas. 'I'd like to add___' 'Are you saying' 'Let me summarize'"	Dialog E vs. A	33% (3)
Teacher revoicing or summarizing student comments	"Re-phrasing what the previous student said to make sure the group understands"	Dialog F vs. A	33% (3)

Materials

The Video Noticing Task and Contrasting Cases Activity were identical to those in Study 1. For the Transcript Annotation condition, we introduced new materials. We created an abbreviated version (cut for length) of Michaels and O'Connor's (2012) *Talk Science Primer* -- a textbook-like chapter that defines and argues for the importance of "academically productive talk," and identifies four goals for productive discussion (e.g., "students listen carefully to one another"). It also outlines nine "talk moves" that teachers can employ to support productive discussion (e.g., asking who can rephrase or repeat; asking for evidence or reasoning). Our abbreviated version was four pages long, printed on standard 8/5 by 11 inch paper. We also provided a two-page transcript of a classroom discussion that had been included in the chapter as an example of productive discussion. As described below, participants in this condition annotated this transcript as part of their learning activity. We removed the original authors' commentary on the transcript.

Procedure

Participants completed activities as part of an introductory, lecture-based class on educational psychology. The course had two sections that met on different days of the week. The study was conducted in both sections. All students were asked to complete activities as part of regular coursework, but were also given the opportunity to volunteer as research participants, allowing us to collect data from their participation. Only those students who consented to be part of the study are included in the analysis.

All participants first completed the pretest Video Noticing Task, but for this study, participants completed this activity at home before class, using Qualtrics software. Protocols for the Video Noticing Task were the same as in Study 1, with the exceptions that information that was presented verbally in Study 1 was presented in written form in Study 2, and for each video clip, participants watched the approximately 90 s clip on one Qualtrics page, and then recorded what they noticed on a new page, being instructed to spend approximately 3 min to write their responses. (A timer counted up to indicate time elapsed, but did not stop participants at 3 min.) The Qualtrics survey was left open for several days, so the pre-noticing task could have been completed between 1 and 4 days prior to the instructional activity.

The instructional activities occurred during class time. Participants were split into experimental condition based on the row in which they were seated: for one side of the class, odd numbered rows (first, third, etc.) completed the Contrasting Cases condition, while even numbered rows completed the Transcript Annotation condition, and for the other side of the class the order was reversed. In this way, half of the first row did each condition, half of the second row completed each condition, and so forth. While not random, this allowed assignment to condition to be highly varied, while still making it easy for students to collaborate with a nearby partner in the same condition.

The Contrasting Cases group completed the Contrasting Cases Activities in pairs, with the same protocols as in Study 1, except that they were not audio recorded. Each participant turned in their own worksheet at the end of the activity. We were not able to consistently collect information about which students worked together as pairs.

The Transcript Annotation group read a four-page expository text about features of good discussion and how to foster it (see Materials, above). They were then instructed to work in

pairs to annotate a single, two-page transcript, underlining and making notes on places they thought indicated discussion features or student or teacher moves that were referred to in the expository text: "As a pair, annotate the transcript to highlight examples of the features of academically productive talk you read about, including the use of talk moves."

After the instructional activities, all participants then listened to the same 15-minute lecture, given by a member of the research team. In order to align well with course content outside the study, the lecture was framed as a discussion of expert noticing more generally, with a focus on teacher noticing during discussion as one context. The lecture presented on the desirable features of classroom discussion and the importance of expert noticing in leading discussions. It then presented on worked examples and contrasting cases as two pedagogical approaches to develop expert noticing. These topics were covered with non-teaching-related examples.

Finally, participants completed the Video Noticing Task a second time, at home in Qualtrics between 0 and 2 days after the instructional experience. The process was identical except that participants who previously watched clips MS1 and HS2 now watched MS2 and HS1, and vice versa.

Coding process

The same senior and junior researchers that coded the data in Study 1 also coded the data in Study 2. For the Video Noticing Task data, we conducted an initial round of coding on a subset of responses to assess the applicability of the coding scheme developed in Study 1 (see Table 1). After some minor refinement to make sure the coding scheme included enough information for coders to categorize responses seen in the undergraduates that were not seen in the PSTs, the remaining responses were all coded by the senior researcher and half were coded by each of the junior researchers (Kohen's kappa average: 0.81 presence, 0.78 absence). All coding was completed blind to condition, as well as to whether a response was given pre- or post-instruction. To create the final data set, the senior researcher used her initial coding results as a base, and compared codes to those of the junior researchers, modifying the base codes as needed based on recognition of an overt coding error or through discussion.

Results

Pre- to post-shifts on video noticing task

Two participants, one each in the contrasting cases condition and the transcript condition, were removed from the analysis for being outliers in the amount of time spent responding to the video noticing questions at either pre or post test (greater than 2.5 SD from the mean, more than 7.7 min average per video).

We then conducted a repeated measures ANCOVA with the within-subject factors of time (pre-post), subject-focus (teacher-focused or student-focused), and valence (presence or absence noted) and between subjects factor of condition (Contrasting Cases or Transcript Annotation), see Fig. 4. Clip order was included as a covariate. There was a significant main effect of time ($F(1,81)=5.82$, $p=.02$, $\eta_p^2=0.07$), but no significant interaction of time by
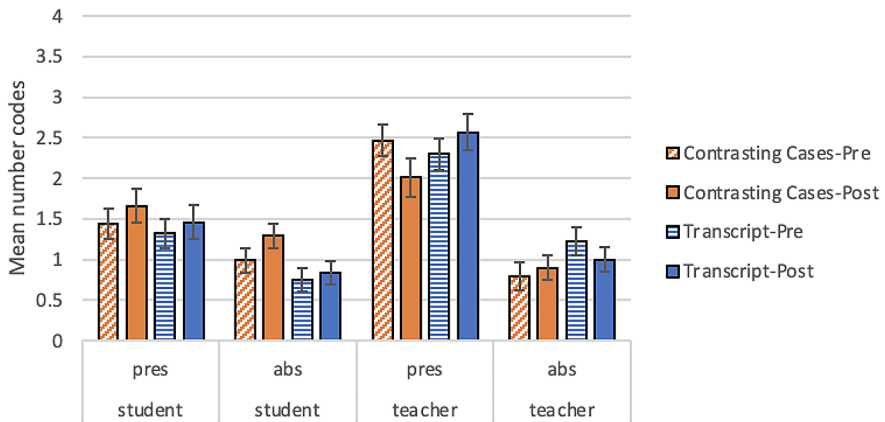
Fig. 4 Mean number of codes by subject-focus (student or teacher) and valence (presence or absence noted) by condition. Error bars represent +/- 1 standard error

condition. There was, however, a small but significant interaction of time by condition by valence (presence vs. absence) ($F(1,81)=4.00$, $p=.049$, $\eta_p^2=0.05$. In other words, change from pre to post in overall number of codes did not differ between conditions, but the conditions differed in patterns in shifts in presence vs. absence codes.

We now look at changes from pre to post-instruction within each condition. For the Contrasting Cases condition, some tests are replications of findings from Study 1, and we adopt directional hypotheses (one tailed tests) in our analysis of these. From pre- to post-test, there was a marginal increase in student-focused codes (t(40)=1.66, $p=.05$, Cohen's $d=0.26$). There was no significant overall change in the number of teacher-focused codes (t(40)=−0.85, $p=.20$). This is similar to the patterns found in Study 1. Given the larger sample size than in Study 1, we look more specifically at the videos of individual teachers to dig into this further (all tests two-tailed, see Fig. 5). The video of teacher MS included more teacher-driven behaviors and less student-led talk and ideas, while the video of teacher HS included more student-led discussion. Examining differences in student-focused codes from pre-to-post for the Contrasting Cases condition, for the video of teacher MS, we see a significant increase in absence student codes (e.g., noting that students did not build on each other's ideas) (t(40)=2.80, $p<.01$, Cohen's $d=0.44$) and an increase in presence student codes for the video of teacher HS (t(40)=2.40, $p=.02$, Cohen's $d=0.37$). Looking at teacher-focused codes, for the video of teacher MS, there was a marginal decrease in noting presence of teacher behaviors (t(40)=-1.80, $p=.08$, Cohen's $d=−0.28$) and a nominal increase in noting absence (or missed opportunities) of teacher behaviors (t(40)=1.50, $p=.15$, Cohen's $d=0.23$). Because of limited sample size, we caution against overinterpreting these findings, but we note that these results are consistent with several of our hypotheses, including an increased focus on student-centered noticing (overall increase in student codes from pre to post), as well as more precise differentiation within student-focused noticing (presence/absence valence, in particular the increase of absence codes for video teacher MS). Evidence of greater precision in differentiating teacher-focused noticing was more
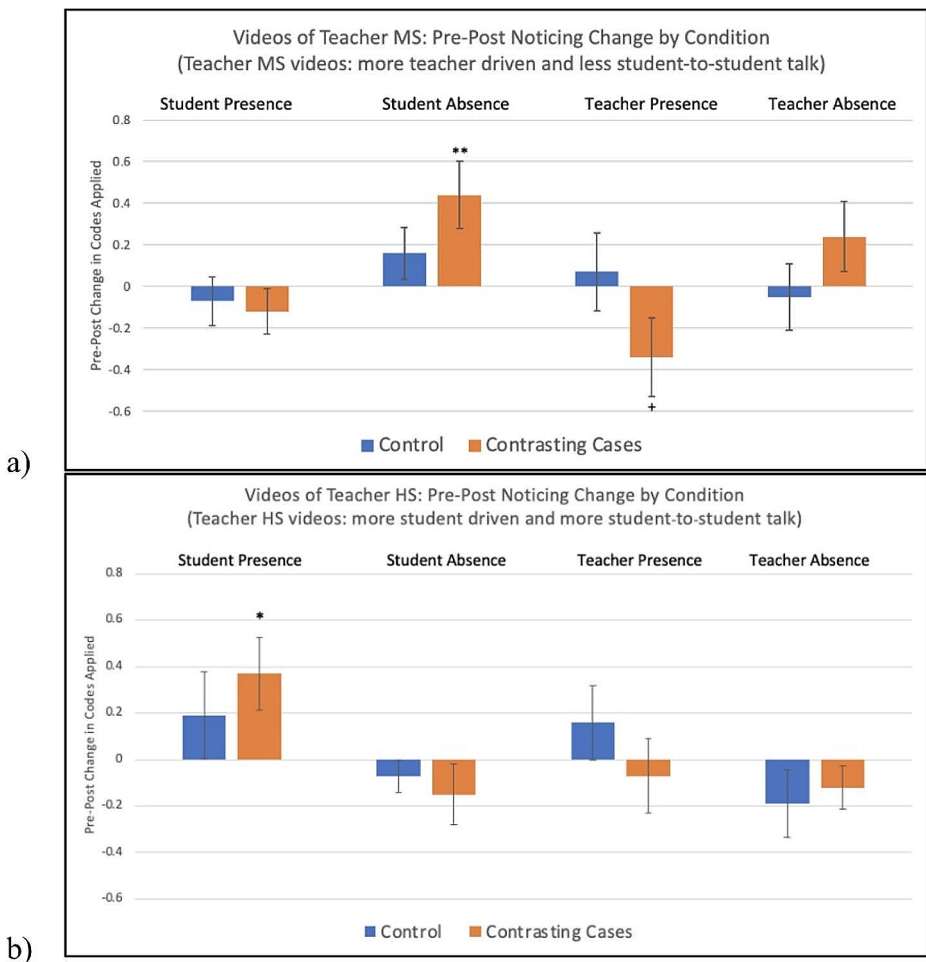
Fig. 5 Pre-Post Differences by Condition (a) for videos of teacher MS and (b) for videos of teacher HS. Error bars represent +/- 1 SE. Stars represent difference from 0 **$p<.01$, * $p<.05$, + $p<.10$

limited, though there was a nominal shift from noting presence to noting more absence for video teacher MS, who ran a very teacher-driven discussion.

For those in the transcript annotation control condition, there were no overall changes from pre to post for either student focused ($t(42)=0.67\ p>.50$) or teacher-focused ($t(42)=0$, $p>.50$) codes. There were also no significant pre-post differences specifically related to either video (all p-values >0.20). One might have expected an increase in presence codes for the videos from teacher HS, as noting the presence of productive talk was the primary activity in the transcript annotation activity. While there was a small nominal increase in student and teacher presence codes for video teacher HS, these were non-significant (student presence: $t(41)=1.0$, $p>.30$; teacher presence: $t(41)=1.1$, $p>.30$).

What happened during the intervention

For participants in the Contrasting Cases condition, we examined their rankings and principles from their in-class worksheets. Due to illegible handwriting or incomplete participant ID information, we were not able to link worksheets for four participants.

To explore how the activity may have influenced noticing, we considered whether the guiding principles generated during the contrasting cases instructional activity predicted pre-post differences in the Video Noticing Task. Six of the most common categories of guiding principles identified by participants directly mapped onto features coded for in the pre-post Video Noticing Task (all categories except the first row of Table 3, "Teacher guiding discussion to keep it on track"). To examine the influence of identifying a principle on shifts in noticing, we created two metrics for each individual who engaged in the contrasting cases activities. One was the average pre to post-instruction Video Noticing Task change on codes for which the participant had identified a corresponding guiding principle during the in-class activity. The second was the average pre-post change on codes for which they had not identified a corresponding guiding principle. Six participants who identified zero of the corresponding guiding principles and one participant who identified all six principles were not included in this analysis. The remaining participants identified an average of 2.03 of the six principles (SD=1.03). A paired subjects t-test found that the average pre-post gain for features that corresponded to an identified guiding principle were higher than the pre-post gain for features for which a corresponding feature was not identified during the activity (*Mean_increase_identified*=0.24, *SD*=0.66 *Mean_increase_non-identified* =-0.01, *SD*=0.43, *t(35)*=2.04, *p*=.049. Note that for each coded feature, the max score is 4, one code for presence and absence for each video clip. The mean score across participants and features was 0.65 at pre.) This finding provides evidence that the process of generating principles through comparison across the cases may have influenced later noticing.

Interestingly, though both the preservice and prospective teachers showed similar patterns in pre-post noticing shifts in Study 1 and Study 2, if we compare the patterns of rankings from the two studies (Figs. 3 and 6), we see a number of differences. Dialog E, which was all student-driven, was universally rated the highest by the preservice teachers in Study 1 (mean ranking 1, SD=0), but showed much more variability in rating among the undergraduate prospective teachers in Study 2 (mean ranking 2.81, SD=1.51). The highest ranked dialog among prospective teachers was Dialog C, which was teacher-led but
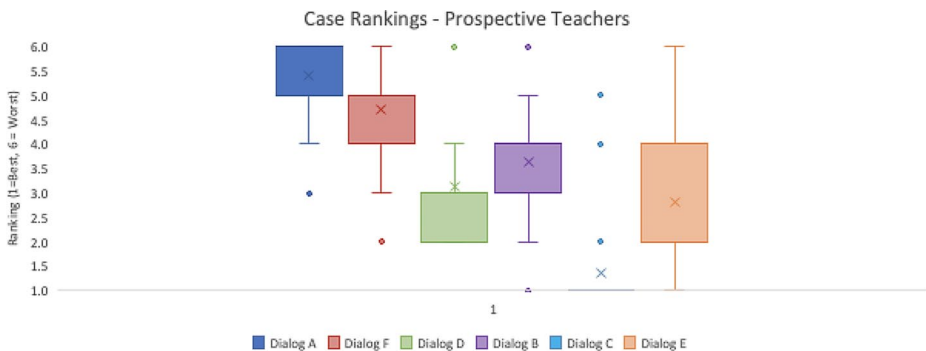
Fig. 6 Ranking of cases by undergraduate prospective teachers (rank 1=best) Dialogs listed in left to right order based on worst to best rank by preservice teachers in Study1

with students contributing most ideas (mean rank=1.34, SD=0.84). Additionally, Dialog D, which included several short response questions from the teacher that were answered by multiple students, was rated more highly by the undergraduate prospective teachers in Study 2 (mean rank=3.13, SD=1.31) than the preservice teachers in Study 1 (mean rank 4.22, SD=0.97) (t(14.94)=2.84, p=.01).

Comparing across Table 2 (Study 1) and Table 3 (Study 2) shows significant overlap in the most common principles identified by participants in the two studies, and these principles were well-aligned with the ideas of uptake that informed design of the dialog cases. Many of the same principles appear in both tables, though in different orders. There were also differences between the two groups. The most common category of principle identified by the undergraduate prospective teachers in Study 2 was "teacher guiding a discussion to keep it on the right track," but this was rarely mentioned by the PSTs (only mentioned by one group) and was not specifically designed for in developing the cases. Conversely, the PSTs in Study 1 were more likely than the undergraduate prospective teachers to develop a principle about student centeredness of discussion (78% of PSTs and 28% of undergraduates) and whether student ideas were validated (44% of PSTs and 16% of undergraduates). Though limited, the student rankings of cases and principles identified provide interesting preliminary information about what different groups of learners value in discussion, and where teacher educators may need to begin when designing learning experiences that shape both knowledge and beliefs surrounding productive class discussion.

Conclusions and significance

Our data show the promise of contrasting cases as a method to improve preservice and prospective teacher noticing relevant to classroom discussion. We posited two ways that changes in noticing might appear based on prior theory: (1) increased precision in recognizing the underlying structure of a phenomenon (e.g., Gibson, 1969), and (2) shifts in focus of attention. We find possible evidence for each in a video-based noticing task that represents transfer from the transcript-based mode of instruction. Effect sizes were modest, but our data support the idea that transcript-based contrasting cases can effectively help tune participants' noticing to be more attentive to students' contributions (beyond a focus on teachers'

Table 3 Most frequent categories of guiding principles in Study 2

Guiding Principle Category	Percent of participants who mentioned (count) N=43
Teacher guiding discussion to keep it on track (e.g., "Teacher as a guide so that [the students] are on the right track. Being directed this way, they don't get off track.")	43% (19)
Teacher restating or summarizing student comments	39% (17)
Teacher asking open-ended or deeper probing questions	36% (16)
Evidence or data brought into discussion	34% (15)
Discussion is student led or student centered	27% (12)
Students building on each other's ideas	27% (12)
Student ideas validated (by teacher or other students)	16% (7)

actions, which is more typical of novice teachers). Increases in student-focused codes were found among both the preservice and prospective teachers in the video noticing task from pre- to post-instruction, compared with a lack of similar overall increase in teacher-focused codes. This pattern may reflect a shift in attention toward students and the features of student talk and ideas. This shift is important given that teachers are more likely to enact in their classrooms the features of discussion that they are able to notice (Barnhart & van Es, 2015). How shifts in noticing translate to actual teaching behavior is beyond the scope of these studies, but one might predict that shifting attention from teacher to students could lead to more advanced discussion methods, such as establishing a classroom culture in which student ideas are valued and students have opportunities to share their thinking even when their ideas differ from others (Huffered-Ackles et al., 2004).

Related to more precise differentiation, increases in noting the absence of features or missed opportunities, found nominally in the Study 1 and significantly in Study 2, may reflect more precision in noticing the structure of student talk. This change entails going beyond merely noticing the existence of talk to noticing the quality of talk -- the source of ideas, true uptake of ideas versus pseudo-uptake, and so forth.

Analysis of discourse data (Study 1) and analysis of the guiding principles dyads created (Studies 1 and 2) both lend support to the role that contrasting cases activities played in bolstering participants' attention to important features of classroom discourse related to what many participants labeled "student centered" discussions. We found shifts in participants' noticing of discussion features to be related to the principles of effective discussion they identified during the activity. Thus, teacher educators may benefit from considering pre-service teachers' beliefs and about effective discussion and ways those beliefs might influence their noticing.

In examining the principles identified and rankings of cases during the activity, we identified differences in what was attended to and identified as good classroom discussion between participants with more versus less teaching experience and exposure to a teacher education program. While our sample was limited, these kinds of findings may enable teacher educators to target learning experiences more specifically along the developmental trajectory of each group of prospective teachers. For example, undergraduate students in education may benefit from opportunities to expand their thinking about the need to control the discussion, from "sage on the stage" to "guide on the side" orientations towards instruction. In contrast, we posit that PSTs, once they have noticed and named teacher moves in the contrasting cases activity, may benefit from opportunities to rehearse and enact the teacher moves in their classrooms and to recognize when they may inadvertently be overcontributing their own ideas and practice holding back. This is in line with prior research that has found that with support and experience, teachers may move towards more student centered dialog (Huffered-Ackles et al., 2004), elicit richer explanations with more equitable distribution of student voices (Banes et al., 2018), and engage more deeply with tensions leading to more purposeful decision making (Sanchez & Athanases, 2023).

Our study has several limitations. The sample size of PSTs was small and limited to those enrolled in a single class within a teacher preparation program. For the undergraduate students, the sample size in each of the experimental conditions was also relatively low, and all students were also from a single university and course. For the undergraduates, we do not have consistent information to identify which students worked together in pairs, prohibiting us from including that information in analysis, such as through multilevel

modeling. Another limitation (though conservative to hypothesis) is imperfect alignment between contrasting cases and the video cases used in the noticing measure. The measures were experimental, and videos were chosen from available online videos to provide an opportunity for noticing in classroom discussions in a more authentic transfer task. They were not developed for the study to specifically address the content covered in the instruction, and as such they do not align perfectly with the features of discussion highlighted in the contrasting cases. Another limitation is that we have not measured whether shifts in noticing in the pre and post-instruction videos were associated with improved abilities to notice while teaching or to lead discussions, which is an avenue for future work. Finally, in constructing the contrasting cases, we made design choices based on the research literature about teachers learning to lead discussions and contrasting cases as an instructional method. However, we did not test alternative possibilities in how the cases could have been designed or what contrasts could have been highlighted. Future research could examine how different features of the cases, including length of dialog, task orientation, discussion features that are contrasted, and effects of video versus written transcripts could influence their effectiveness.

We have demonstrated that a contrasting cases approach to noticing class discussion may offer benefits over the more commonly used approach of single transcript analysis, and may support novice teachers in their trajectory as orchestrators of effective class discussion. Despite the benefits of classroom discussion as a pedagogy (Mercer & Hodgkinson, 2008; Juzwik et al., 2013), many novice and pre-service teachers lack substantive experience participating in, much less facilitating good discussions (Kavanagh et al., 2019). Even when given the opportunity to observe discussions, new teachers need to learn to notice the important features of such learning environments. Classroom talk is fleeting. Subtle features, such as what happens to a student contribution or how various ideas are connected, can be difficult to attend to in the moment with many competing demands for teacher attention. As such, teacher development that supports *perceptual skills* in noticing may support capacity to make informed choices in practice (Yanow & Tsoukas, 2009). Our results highlight the potential of contrasting cases to support such development, and we have tried to describe our design principles in ways that could be generative for future research and design. For example, future work might consider additional dimensions of productive discussion and attend more specifically to issues of equity and support for culturally and linguistically diverse learners.

Appendix: contrasting cases dialogs and activity

Read through each of the following excerpts. They represent different potential dialogs from one small part of a class discussion. In them, students are using articles, maps, and graphs to decide whether or not the placement of toxic waste sites is an environmental justice issue. The goal is that across them, these short, targeted excerpts will highlight different features of how students and teachers are picking up and building on each other's ideas.

A) As a pair, rank the 6 dialogs in order from what you consider the least good example of a discussion to the best example. There is no right answer here.
B) Think about what made you choose the ranking you did, focusing particularly on the taking up and building of ideas. As a pair, identify 2–3 principles that guided how

you ranked them. Taken together, your principles should explain the ranking of all the dialogs (e.g., for any dialog pair, someone should be able to look at your principles to decide which one you likely thought was better.) You should be able to imagine applying these principles to a new video or transcript of a discussion.

1)
  2)
  3)
  C) What was your ranking of the six dialogs? Write the letters (A-F) in order from least good to best.

Least good Best.

Dialog A

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

T: Okay. Good! Someone on the other side of the room want to share? Can you add something new?

S2: I'd like to add that rich people don't get sick because they don't live by the toxic waste sites. They're just mostly by low-income neighborhoods.

T: Great! That's one perspective. Does anybody see it differently?

S3: Well, people could just move away from the toxic waste if they don't like it.

Dialog B

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

T: So, what I hear you saying is that the location makes it an environmental justice issue because they don't make people sick evenly, right? So, not everyone is impacted equally by the toxic waste sites. Someone on the other side of the room want to share? Can you add something new?

S2: I'd like to add that rich people don't get sick because they don't live by the toxic waste sites. They're just mostly by low-income neighborhoods.

T: Right! So, the map is showing us that toxic waste sites are usually located near low-income communities, where the article said people are more likely to have asthma and other health problems. David and Jose, it sounds like you are both saying that makes it an environmental justice issue. Does anybody see it differently?

S3: Well, people could just move away from the toxic waste if they don't like it.

Dialog C

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

T: So, what I hear you saying is that the location of toxic waste sites is an environmental justice issue because they make people sick. Can you tell me more about why you think it's a justice issue in particular?

S1: They don't make people sick evenly. Like, white people and people with money don't get sick as much.

T: Yeah. So not everyone is affected equally by the toxic waste. Someone on the other side of the room want to share? Can someone add to that using the data?

S2: I'd like to add that rich people don't get sick because they don't live by the toxic waste sites. They're just mostly by low-income neighborhoods.

T: How did you figure that out? Can you tell us what you saw in the map that told you that?

S2: I saw that the area around the toxic waste sites is red on the map. The red means that it's a low-income neighborhood, and the pink and yellow are more expensive houses. But the area around the toxic waste sites is red in almost every area of the map.

T: Does anybody see it differently?

S3: Well, people could just move away from the toxic waste if they don't like it.

Dialog D

SI: talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

T: So, what I hear you saying is that the location of toxic waste sites is an environmental justice issue because they make people sick. Who gets sick more and who gets sick less….

S1: Like, white people and people with money don't get sick as much.

T: Yeah. So, not everyone is affected equally by the toxic waste. Someone on the other side of the room want to add? Maybe about what kinds of neighborhoods…Toxic waste sites are usually near….

S2: Near low-income neighborhoods.

T: Right. What told you that from the map? The area around the toxic waste sites is what color?

S2: Red.

T: Which means….

S2: Low income.

T: Good. So we see from the map that low income neighborhoods are more affected, which makes it unequal and a social justice issue. Anyone see it differently?

S3: Well, people could just move away from the toxic waste if they don't like it.

Dialog E

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

S2: Can you say more about what you mean by that?

S1: Like, they don't make people sick evenly. White people and people with money don't get sick as much.

S3: I'd like to add to David because rich people don't live by the toxic waste site. The toxic waste sites are mostly near low-income neighborhoods.

S4: So, are you saying that the red on the map means that it's a low-income neighborhood?

S3: Yeah, and the area around the toxic waste sites is red in almost every area of the map.

S1: Ok. Let me try to summarize. The map is showing us that toxic waste sites are usually near low-income neighborhoods, and cause low income people more health problems. And that's why we think it's an environmental justice issue. Does anybody think something different?

S3: Well, people could just move away from the toxic waste if they don't like it.

Dialog F

S1: I talked about how toxic waste sites are kinda like the lead paint we read about last week—they both make people sick.

T: Okay. Good! So toxic waste makes people sick. Anyone else want to add to that?

S2: I'd like to add that rich people don't get sick because they don't live by the toxic waste sites. They're just mostly by low-income neighborhoods.

T: Alright. You're saying rich people don't get as sick, and the waste sites are by low income neighborhoods. Does anybody else see it differently?

S3: Well, people could just move away from the toxic waste if they don't like it.

SUPPLEMENTAL CASES INTRODUCED AFTER COMPLETING THE FIRST ACTIVITY

Do your principles apply to raking these short dialogs? You can add to them or refine them your principles you'd like.

Dialog G

S1: Like White people and people with money don't get sick as much from the toxic waste. So that means it's not fair.

S2: I agree with David because rich people don't live by the toxic waste sites. The toxic waste sites are mostly near low-income neighborhoods like we can see on the map.

Dialog H

S1: Like White people and people with money don't get sick as much from the toxic waste. So that means it's not fair.

S2: I agree with David because pollution is really gross, like when there's garbage in the ocean and it kills the dolphins. People shouldn't pollute the earth.

Dialog I

S1: Like White people and people with money don't get sick as much from the toxic waste. So that means it's not fair.

S2: People shouldn't get punished and get sick more because of where they live. They can't afford to move.

Appendix B

Matrix of dialog feature variations. Each row represents a feature and how it was across the dialogs.

	Dialog A	Dialog B	Dialog C	Dialog D	Dialog E	Dialog F
Engaging multiple student voices	3 students	3 students	2 students	3 students	4 students	3 students
Eliciting and probing student ideas	Open question, no probing	Open question, no probing	Open question with teacher probing	Leading short answer questions+probes	Open question with student probing	Open question, no probing
Revoicing	No	Yes- teacher injects ideas	Yes - teacher	Yes - teacher	Yes – student	Yes - teacher
Uptake and building of student ideas	Low	Yes - teacher uptake	Yes – student uptake	Teacher directed student uptake	Yes – student uptake	Low
Bring evidence into explanations	No	Teacher provides evidence	Students provide evidence	Students provide teacher directed evidence	Students provide evidence	No

.

References

Adler, M., Rougle, E., Kaiser, E., & Caughlan, S. (2003/2004). Closing the gap between concept and practice: Toward more dialogic discussion in the language arts classroom. *Journal of Adolescent & Adult Literacy*, *47*(4), 312–322.

Athanases, S. Z., & Sanchez, S. L. (2019). Challenges for preservice teachers in learning to lead discussions: Inquiry in diverse classrooms. *Paper presented at the annual meeting of the American Educational Research Association (AERA)*, Toronto, ON, Canada.

Banes, L. C., Ambrose, R. C., Bayley, R., Restani, R. M., & Martin, H. A. (2018). Mathematical classroom discussion as an equitable practice: Effects on elementary English learners' performance. *Journal of Language Identity & Education*, *17*(6), 416–433.

Banes, L. C., Restani, R. M., Ambrose, R. C., Martin, H. A., & Bayley, R. (2020). Relating performance on written assessments to features of mathematics discussion. *International Journal of Science and Mathematics Education*, *18*, 1375–1398.

Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among pre-service science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*, *45*, 83–93. https://doi.org/10.1016/j.tate.2014.09.005.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* (Vol. 24, pp. 61–100). https://doi.org/10.3102/0091732X024001061.

Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: Inventing a general solution versus compare and contrast. *Instructional Science*, *44*, 177–195.

Cirillo, M. (2013). What does research say the benefits of discussion in mathematics class are? *National Association of Teachers of Mathematics*, (Research Brief No. 19).

Davis, B. G. (1993). *Tools for teaching*. Jossey-Bass Inc.

Davis, E. (2006). Characterizing productive reflection among preservice elementary teachers: Seeing what matters. *Teaching and Teacher Education*, *22*(3), 281–301.

Derry, M. J., Wilsman, Alan, J., & Hackbarth (2007). Using contrasting Case activities to deepen teacher understanding of Algebraic thinking and teaching. *Mathematical Thinking and Learning*, *9*(3), 305–329. https://doi.org/10.1080/10986060701361033.

Erickson, F. (2011). On noticing teacher noticing. In M. Sherin, V. Jacobs, & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 17–34). Routledge. https://doi.org/10.1016/j.tate.2005.11.005.

Forzani, F. M. (2014). Understanding core practices and practice-based teacher education: Learning from the past. *Journal of Teacher Education*, *65*(4), 357–368. https://doi.org/10.1177/0022487114533800.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Meredith Corporation.

Herbel-Eisenmann, B., Drake, C., & Cirillo, M. (2009). Muddying the clear waters: Teachers' take-up of the linguistic idea of revoicing. *Teaching and Teacher Education*, *25*(2), 268–277. https://doi.org/10.1016/j.tate.2008.07.004.

Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education*, 81–116. https://doi.org/10.2307/30034933.

Juzwik, M. M., Borsheim-Black, C., Caughlan, S., Heintz, A., & Nystrand, M. (2013). Inspiring dialogue: Talking tolearn in the English classroom. Teachers College.

Kavanagh, S. S., Monte-Sano, C., Reisman, A., Fogo, B., McGrew, S., & Cipparone, P. (2019). Teaching content in practice: Investigating rehearsals of social studies discussions. *Teaching and Teacher Education*, *86*, 1–11. https://doi.org/10.1016/j.tate.2019.06.017.

Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, *68*(5), 845–861. https://doi.org/10.1177/0013164407313369.

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*(3), 568–589. https://doi.org/10.3102/0002831212437853.

Kisa, M. T. (2013). *Science teachers' learning to notice from video cases of the enactment of cognitively demanding instructional tasks* [Doctoral dissertation, University of Pittsburgh]. ResearchGate.

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education*, *60*(2), 142–154. https://doi.org/10.1177/0022487108330245.

Lin-Siegler, X., Shaenfield, D., & Elder, A. D. (2015). Contrasting case instruction can improve self-assessment of writing. *Educational Technology Research and Development*, *63*(4), 517–537. https://doi.org/10.1007/s11423-015-9390-9.

Mehan, H. (1982). The structure of classroom events and their consequences for student performance. In P. Gillmore & A. Glatthorn (Eds.), *Children in and out of school* (pp. 59–87). Washington, DC: Center for Applied Linguistics.

Mercer, N., & Hodgkinson, S. (Eds.). (2008). *Exploring talk in school: Inspired by the work of Douglas Barnes*. Sage.

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, *30*(3), 359–377. https://doi.org/10.1080/01411920410001689689.

Michaels, S., & O'Connor, C. (2012). Talk science primer. *TERC*.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussions. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing Intelligence through Talk and Dialogue* (pp. 333–347). Washington DC: AERA. https://doi.org/10.3102/978-0-935302-43-1_27.

Murphy, P. K., Wilkinson, I. A., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*(3), 740. https://doi.org/10.1037/a0015576.

Nagarajan, A., & Hmelo-Silver, C. (2006, June). Scaffolding learning from contrasting video cases. *In Proceedings of the 7th International Conference on Learning Sciences: Indiana University. Volume 1* (pp. 495–501). International Society of the Learning Sciences. https://doi.org/10.22318/icls2006.495.

Nystrand, M. (1997). *Open dialogue: Understanding the dynamics of language and learning in English classrooms*. Teachers College.

O'Connor, C., & Michaels, S. (2019). Supporting teachers in taking up productive talk moves: The long road to professional learning at scale. *International Journal of Educational Research*, *97*, 166–175. https://doi.org/10.1016/j.ijer.201\\7.11.003.

Pianta, R., & Hamre, B. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119. https://doi.org/10.3102/0013189X09332374.

Reznitskaya, A. (2012). Dialogic teaching: Rethinking language use during literature discussions. *The Reading Teacher*, *65*(7), 446–456. https://doi.org/10.1002/TRTR.01066.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*, 561–574. https://doi.org/10.1037/0022-0663.99.3.561.

Roelle, J., & Berthold, K. (2015). Effects of comparing contrasting cases on learning from subsequent explanations. *Cognition and Instruction*, *33*(3), 199–225. https://doi.org/10.1080/07370008.2015.1063636.

Sanchez, S. L., & Athanases, S. Z. (2023). Dialogic teacher inquiry: The case of a preservice teacher learning to facilitate class discussion. *Dialogic Pedagogy: An International Online Journal*, *11*(1), A1–A38. https://doi.org/10.5195/dpj.2023.482.

Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM-Mathematics Education*, *43*(1), 133–145. https://doi.org/10.1007/s11858-010-0292-3.

Schenke, K., & Richland, L. E. (2017). Preservice teachers' use of contrasting cases in mathematics instruction. *Instructional Science*, *45*(3), 311–329. https://doi.org/10.1007/s11251-017-9408-2.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–5223.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*, 129–184. https://doi.org/10.1207/s1532690xci2202_1.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*, 759–775. https://doi.org/10.1037/a0025140.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. WW Norton & Company.

Shemwell, J. T., Chase, C. C., & Schwartz, D. L. (2015). Seeking the general explanation: A test of inductive activities for learning and transfer. *Journal of Research in Science Teaching*, *52*(1), 58–83.

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, *11*(2), 107–125.

TeachingWorks (2023, January 26). *High leverage practices*. https://www.teachingworks.org/high-leverage-practices/.

van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM–Mathematics Education*, *53*(1), 17–27. https://doi.org/10.1007/s11858-020-01211-4.

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, *78*(3), 516–551. https://doi.org/10.3102/0034654308320292.

Wiens, P. D., Hessberg, K., LoCasale-Crouch, J., & DeCoster, J. (2013). Using a standardized video-based assessment in a university teacher education program to examine preservice teachers knowledge related to effective teaching. *Teaching and Teacher Education*, *33*, 24–33.

Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, *28*(2), 193–219. https://doi.org/10.1177/0741088311399236.

Yanow, D., & Tsoukas, H. (2009). What is reflection-in-action? A phenomenological account. *Journal of Management Studies*, *46*(8), 1339–1364. https://doi.org/10.1111/j.1467-6486.2009.00859.x.