Video analysis for insight and coding: Examples from tutorials in introductory physics

Rachel E. Scherr

Physics Education Research Group, Department of Physics, University of Maryland, College Park, Maryland 20742, USA
(Received 17 April 2009; published 19 October 2009)

The increasing ease of video recording offers new opportunities to create richly detailed records of classroom activities. These recordings, in turn, call for research methodologies that balance generalizability with interpretive validity. This paper shares methodology for two practices of video analysis: (1) gaining insight into specific brief classroom episodes and (2) developing and applying a systematic observational protocol for a relatively large corpus of video data. These two aspects of analytic practice are illustrated in the context of a particular research interest but are intended to serve as general suggestions.

I. INTRODUCTION

The increasing ease of video recording offers physics education researchers new opportunities to create richly detailed records of classroom activities. The level of detail can be as daunting for analysis as it is inspiring. In a few weeks or months, a researcher can generate dozens or even hundreds of hours of videotape. The Physics Education Research Group at the University of Maryland, for example, has in the past seven years accumulated approximately 2000 hours of video of small-group collaborative active-learning activities ("tutorials") in introductory university physics courses. The question of how to find what you are looking for—or even what to look for—can quickly become overwhelming.

This paper shares methodology for two practices of video analysis: (1) gaining insight into specific brief classroom episodes and (2) developing and applying a systematic observational protocol for a relatively large corpus of video data. These two aspects of analytic practice are illustrated in the context of a particular research interest of mine—the dynamics among behavior, epistemological framing, and the conceptual substance of student reasoning[1]—but the analysis methodologies are intended to reach beyond this particular interest.

II. THEORETICAL FRAMEWORK

No methodology is without its underlying theoretical premises. The practice of using rich records of naturally occurring activities as evidence of student knowledge promotes and supports a particular point of view: that learning and expertise show best in what students do and say to learn together. This view contrasts with a perspective in which student knowledge is best represented with private mental activities and the primary evidence of learning and expertise are in the readings of specialized probes (such as specially designed written questions).[2] Because video shows the ways in which knowledge practices are situated in the interactions of people with each other and with material objects, perspectives from actor-network theory,[3–5] embodied cognition,[6] situated or distributed cognition,[7,8] and activity theory[9] are naturally relevant.

III. DATA COLLECTION

I describe general features of the data collection that has taken place at the University of Maryland in 2001–2008 in order to give a sense of what is involved in our case.

A. Instructional setting

Tutorials are a particularly rich instructional setting for observing what students do to learn together. Because they are working in collaborative groups, we have the opportunity to watch them talk and interact, both among themselves and with instructors. They work in groups of three or four, which is large enough to support diverse interactional patterns but small enough so that most of the students participate. The students mostly sit still at tables, which facilitates relatively unobtrusive videotaping of single groups for the duration of a class.

The tutorials that will serve as examples for the analysis methodology that follows took place as part of a two-semester algebra-based introductory physics course at the University of Maryland, with approximately 160 students in each lecture section, most of whom are junior and senior health and life science majors. More than half are female and there is wide ethnic diversity reflecting the student population of the University of Maryland. The course was reformed as part of a project titled *Learning How to Learn Science: Physics for Bioscience Majors*, which was carried out at the University of Maryland from 2000 to 2005.[10]

Tutorials at the University of Maryland are, like those developed at the University of Washington, worksheet-based group-learning activities.[11,12] In the tutorial sessions, students work in small groups on sequences of questions that lead them to make predictions and compare various lines of reasoning in order to build an understanding of basic concepts. Teaching assistants (TAs) serve as facilitators rather than as lecturers. (Observational studies suggest that this simple idea does little to convey the complexity of TAs' actual classroom practice.[13–15]) Each class section consists of five groups of four students each, supervised by two TAs. University of Maryland tutorials are constructed to emphasize the reconciliation of everyday intuitive thinking and experience with formal scientific thinking, as well as to encourage explicit epistemological discussions about the learning process.[16] They are distributed as part of a project whose purpose is to provide instructors with the materials and resources to implement similar tutorials at their own institutions (titled *Helping Students Learn How To Learn: Open-Source Physics Work-*

sheets Integrated With TA Development Resources).[17]

B. Goals of data collection

Projects at the University of Maryland have collected video data for a variety of purposes. The *Learning How To Learn Science* project mentioned above used tutorial video to seek examples of student epistemological resources and metacognitive activity during instruction. The *Helping Students Learn How To Learn* project and its sequel, *Open-Source Tutorials*, made extensive use of video to assess the effectiveness of the tutorials and to develop resources for instructors who wanted to implement them. A third project, *Developing Conceptual and Teaching Expertise in Physics Graduate Students*, has specifically followed particular TAs in their classroom interactions with students. All together, these projects have accumulated approximately 2000 hours of video over about seven years. The video is mostly of tutorials but also includes some laboratories, lectures, homework sessions, and interviews.

While the video was in some cases collected for specific purposes, it is rich enough to support analysis from a wide variety of perspectives. Our video library is a valuable resource for ongoing projects.

C. Method of data collection

The nature of the activity in which students are engaged (as perceived by the students) affects their interactions.[6,18–21] Thus, we try to keep the recording of the tutorial activities subordinate to normal classroom practices. At the University of Maryland, tutorials are held in a single room with six tables at which students work collaboratively in groups of four; students typically do not move from their seats during the class session or even from week to week. For video recording, two small video cameras on tripods are positioned on the periphery of the room, each focused on a single table. The cameras do not move. Microphones are embedded in cages on the tables that are being recorded. A researcher turns on the cameras at the start of the tutorial session, but the cameras are otherwise unattended. Our intention is to make the video recording as unobtrusive as is practical in the available physical environment even at the expense of visual or sound quality. (If we were designing a video-equipped classroom to be built to our specifications, we would plan for different recording practices.)

Any research needs to take into account the extent to which data collection may have affected the events being analyzed, and the question often arises of whether students are influenced by the presence of the camera. I take the position that the question of camera effects is an empirical question to be decided on each occasion.[2] Participants tend to demonstrate when they are attending to the camera by interacting with the recording equipment or changing their behavior in an observable way. Mostly, we have found that people habituate to the camera quickly if they notice it at all. This result is consistent with other research on camera effects.[22] In the case of tutorials, exceptions seem to occur when participants engage in an activity that they recognize as inappropriate: after someone uses profanity, breaks experimental equipment, or jokes with a peer, participants sometimes turn their gaze to the camera, address the microphone, or say something about getting caught. This orientation to the recording equipment can be a useful marker for activities that the participants understand to be improper.

IV. DATA ANALYSIS

My own analysis of video data takes two major forms: *insight oriented* and *coding oriented*. The primary purpose of insight-oriented analysis is to identify and richly describe the events in the episode as they unfold in the interaction among the participants. Coding-oriented analysis, in contrast, aims to generate categories that will be used to document the frequency and distribution of events of interest. Of course, insight-oriented analysis may generate a coding scheme, and coding may lead to deeper insights. In what follows, both forms of analysis are illustrated with examples from research on the dynamics among behavior, epistemological framing, and the conceptual substance of student reasoning. These examples are excerpted from previously published work.[1]

A. Gaining insight into specific episodes

As conducted in our research group, insight-oriented analysis begins with episode selection and is followed by collaborative viewing. Episode selection is typically accomplished by an individual researcher. Collaborative analysis is ideally conducted by a group of researchers with diverse backgrounds and interests and is the primary means for reaching detailed evidence-based conclusions about short episodes.

1. Selecting episodes

In the rich medium of video, selection criteria can be elusive. When searching a large video library for useful episodes, I find that my research priorities are sometimes revealed in the way I search. For example, I once observed that I tended to look for episodes in which the students are engaged in high-quality scientific reasoning. It took some effort on my part to articulate this priority: to myself, I had just said I was looking for "good stuff." An orientation to high-quality scientific reasoning, while natural for a physics education researcher, is by no means the only possible predilection: other analysts might look for incorrect student ideas, moments in which the tutorial worksheet causes trouble for students, "aha" moments, patterns in who initiates interactions between students and instructors,[23] or any number of other kinds of events or patterns.

In looking for my particular kind of good stuff, I usually began by sitting alone with a video camera using the camera's small display to scan the tape. Without an explicit research interest to focus my attention, I found myself half-unconsciously fast-forwarding the videotape to skip the "boring" parts. When I spotted the students sitting up straight and gesticulating at one another, I stopped the tape there to hear what they were saying. Something in me apparently thought that those moments were more likely to include good stuff than the periods in which students were bent over their

worksheets, writing privately or occasionally exchanging information. Once I realized what I was doing, I stopped and decided to make my unconscious assumption the subject of a systematic analysis: can we tell something about the quality of students' reasoning just by looking at them?

There was reason to think that students who were sitting up, speaking clearly, and gesturing frequently were in fact engaged in a different kind of activity than those who were talking in subdued voices with their hands relatively still. *Framing* is a construct developed in anthropology and linguistics[24] to describe how an individual or group forms a sense of "what it is that is going on here."[18–21] Previous analyses have found evidence of framing in the substance of participants' speech and in linguistic markers associated with speech acts.[25] However, because frames "emerge in and are constituted by verbal and nonverbal interactions,"[20] paraverbal and nonverbal cues may also present evidence of framing. In any social interaction, framing presents a communicative task in which participants collaboratively establish the nature of their shared activity.[20] Conversational partners mutually construct a type of activity not only by speaking and listening to one another but also by performing and observing body language (e.g., facial expression, gesture, and posture), prosodic features of utterances (e.g., pitch variation, loudness, pausing, and pacing), and linguistic signals (choice of vocabulary, level of formality, and choice of pronouns).[6,20] Verbal and nonverbal displays often reinforce one another, which we expect contributes to the coherence of participants' framing, and we may hope to identify their framing on the basis of those displays. For example, if a speaker makes eye contact, gestures while speaking, and uses an animated voice, she is likely both experiencing an engaged discussion herself and displaying to others that they are mutually engaged in a discussion.[26] If the content of her speech reinforces the show of engagement (perhaps by its novelty, emotional content, intellectual insight, or connection to experience), then we have multiple strands of evidence that she, as an individual, is framing the activity as an engaged discussion. If her partners respond in kind, both verbally and nonverbally, there is evidence that the other participants share the framing as well.

My unconscious impulses aligned well enough with the research literature that I set out to formulate a systematic observational protocol. I selected episodes that I thought showed students making a transition from one tutorial framing to another and brought them to my research group for collaborative analysis.

2. Analyzing episodes collaboratively

The backbone of video analysis as we conduct it in the Physics Education Research Group is collaborative analysis of short episodes of classroom activity. Usually, one person (the "owner" of the data) brings preselected episodes in which he or she has an interest. These episodes are usually short; we often find that five minutes of video data contain enough activity for an hour of discussion. The owner plays the video, and members of the group call to stop the playback whenever they see something they want to discuss.[27] We often benefit from replaying the video multiple times. In many cases, we do not approach the data with preconceived categories of interest. Rather, the goal of the group work is to progressively deepen our understanding of the participants' activities and to challenge the biases of individual analysts. Topics for discussion are unlimited: physics content, vocal tone, instructional moves, nonverbal behavior, and any other themes of interest to the participants are invited. There is, however, a constant effort to ground assertions in the evidence of the video episode and to limit proposed hypotheses to those for which the video episode (or others like it) could provide confirming or disconfirming evidence. The resulting discussion is often a lively multidimensional conversation reflecting the diverse interests of the participants as well as the focal interest of the data's owner.

3. Examples of detailed analysis

The first analysis of a short segment of video below provides a brief illustration of the kind of detailed observations that can result from discussion of a segment only 14 seconds long. Following that analysis is an analysis of a two-minute episode of a different group working on the same tutorial.[28]

(a) *A segment of a few seconds.* In an excerpt from students working on a tutorial on Newton's third law (video 1), Jasmin[29] says:

> *In summary, for most people* blah blah blah blah… *We shouldn't dwell on this kind of question and instead focus on learning exactly when…* This is like one of the questions.

The italicized portion of the statement consists of abbreviated but otherwise verbatim statements from the tutorial worksheet. While she says those words, Jasmin is leaning over and looking at the worksheet, using a singsong tone, speaking softly, using nonwords ("blah blah"), and hugging her arms closely to her body. Her behavior matches the evidence from her speech, which is that she frames the activity as a form of *reading aloud from a common text.* By both her physical behavior and the content of her speech, Jasmin indicates specific expectations for this activity: when you read, you look at what you are reading. A singsong tone and lack of gestures signal recitation.[30,31] If one's peers are reading the same text, it is appropriate to speak quietly and abbreviate, since the point is more to indicate what you are reading than to ensure others' comprehension.

Then, when Jasmin says, "This is like one of the questions," she speaks more loudly without the singsong recital tone, pushes back from the table, and points at her paper with her pencil. These paraverbal and nonverbal behaviors help her peers to recognize that she is doing something different now: Jasmin frames the new activity as *commenting on what she has just read* (in particular, pointing out a similarity between the worksheet task and another question she expects is familiar to her peers). Again, her behavior matches the substance of what she is saying. Her clear loud voice denotes original speech rather than recitation; she indicates that she is now talking *about* the worksheet (rather than reading from it) both by pointing to it and by distancing her body from it. One of her partners, Sheryl, recognizes this change in the moment and knows what she is referring to: she sits up,

looks at Jasmin, and says, "Yeah, this is question six on the homework."

Our specific observations of this very short segment of video and others like it have informed the development of general theoretical principles in our research group. Humans are continually confronted with interpretive tasks such as the one Sheryl faced in understanding Jasmin. In order to understand what an utterance means, we have to sense what kind of activity the utterance is part of; that is, we have to frame the activity. Is it reading, commenting, explaining, imitating, or chatting? We make this interpretation based on verbal, paraverbal, and nonverbal cues in the interaction—based not only on *what people say* but also on *how they say it*—and we think and act in response to the interpretation we make. That is, nonverbal and paraverbal behavior provide both participants and researchers with evidence of framing.

(b) *An episode of a few minutes.* The next example shows how we began to analyze student interactions with an eye to how verbal and nonverbal displays reinforce one another, contributing to the participants' mutual understanding of what is taking place and providing evidence for research. Our insights at this point were beginning to suggest an observational methodology that we might use to explore dynamics among behavior, framing, and the conceptual substance of student reasoning in the context of tutorials. In addition, we had begun to accumulate evidence that certain student behaviors indicate and support a more sophisticated epistemological framing of tutorial activities, one in which students discuss the substance of the ideas at hand.

The students in this 2.5 minute episode (video 2) are also working on the Newton's third law tutorial (in the same room as Jasmin's group during the same hour). The tutorial worksheet begins by stating Newton's third law and admitting that in some cases it seems not to make sense (an admission that is well supported by research into student understanding of Newtonian mechanics).[32–34] The worksheet asks students to consider a heavy truck ramming into a parked unoccupied car: "According to common sense, which force (if either) is larger during the collision: the force exerted by the truck on the car or the force exerted by the car on the truck?" Many students believe that the force by the truck is larger. In order to help students reconcile the contradiction between their intuition and Newton's third law, the tutorial helps them consider the acceleration of each vehicle during the collision.

In the part of the tutorial represented below, students are asked to imagine that the truck weighs twice as much as the car and slows down by 5 m/s during the collision. In this episode, Luisa (L), Erin (E), Yvonne (Y), and Hannah (H) respond to the question: "Intuitively, how much speed does the car gain during the collision?" The exercise is reproduced in Fig. 1. A detailed description of the tutorial is in Ref. 16.

Yvonne initially proposes that the car should gain the same amount of speed as the truck loses. Luisa suggests that the car should gain twice as much speed as the truck loses since it has half the mass of he truck, and Hannah agrees with her. Luisa and Erin comment on the possible intentions of the authors of the worksheet. In the last part of the episode, a word overheard from the next table seems to motivate the students to clarify what they think is going on in the collision.

III. A new strategy: Refining intuition

Before accepting that there's an irreconcilable contradiction between Newton's third law and the intuition that the car reacts more during the collision, let's try a reconciliation strategy called *refining your intuitions.*

A. We'll start with a new question. Suppose the truck's mass is 2000 kg while the car's mass is 1000 kg. And suppose the truck slows down by 5 m/s during the collision. Intuitively, how much speed does the car gain during the collision? (Apply the intuition that the car reacts more during the collision, keeping in mind that the truck is twice as heavy.) Explain your intuitive reasoning.

FIG. 1. Excerpt of tutorial on Newton's third law.

H: *Intuitively, how much speed does the car gain*
Y: Does it gain 5 m/s?
H: It should, just because
Y: Truck slows down 5 m, car speeds up 5 m/s (laughs)? I don't know
H: *Apply the intuition that the car*
L: Shouldn't it gain ten though?
E: Or gain twice as
L: Right, because the car is
H (with L): Yeah, because it's twice as heavy
L: Right.
H: So, it should
L: The car is half as heavy, so it'll gain twice as much
Y: Ah, shoot (laughs)
H: Or something, I don't know
Y: Wow, okay.
H: Okay, when
L: That's what they want us to think, but this is not the real answer.
E: This is not the right one, this is not the right one. Apparently, I think that's what they want us to say.
H: This is going…5… 5 m/s, that's it's what, acceleration or velocity?
Y: Speed. Velocity.
L (together with Y): Velocity.
H: Velocity.
E: *Slows down by*
Y: So the car gains 10 m/s?
E: I guess.
H: Didn't he say something about how like— somebody in class, like—if something's touched a velocity or something was changed… what was he talking about in class, something, the masses?
[TA at next table: So what would you think intuitively?]
[Student at next table: Intuitively I was thinking that

some (inaudible)]
H: I hate that word, intuitively.

Y: See intuitively, I would think that it'd slow down, I mean speed up 5 m/s.
E: Isn't the car—do they mean once they hit?
H: Because it is slowing down.
Y: If the if the truck if the truck if the truck, suppose the truck slows down by 5 m in the collision, so if the truck is slowing down, then the car has to be speeding up.
H: The car's not moving!
L: Yeah, it's not moving.
E: But I think does that mean once it hits
Y: No, but I'm saying it says how much does the car gain, cause it, yeah
E: Cause like it goes (gestures collision)
Y: Yeah, that's, yeah.
E: Now wait, would actually…
Y: I think it would go from zero to five, instead of, you know, being still.
H: All right, if you think about it them being the same weight… (gestures)
L: Yeah
H: then do you think that it would push it like the same
Y: The same.
L: Then I would think it would push it the same.

During the first 1.5 minutes of the episode, students lean over their papers and only briefly glance up at their peers. They rarely gesture, and their faces and voices are relatively neutral. Three of the students regularly rest their foreheads on one hand. While their physical behaviors are of this type, the substance of their interactions focuses on completing the worksheet and they offer only brief answers with minimal reasoning for support. Hannah asks for verification of the technical label for 5 m/s ("That's it's what, acceleration or velocity?"), and Yvonne and Luisa respond factually. Yvonne tries to establish their response to the question ("So the car gains 10 m/s?") and Hannah tries unsuccessfully to recall something said in class about how velocities may be changed. Some of their statements explicitly mark the possibility that their answers are not genuine: Luisa says "That's what they want us to think, but this is not the real answer," and Erin agrees that "This is not the right one… Apparently, I think that's what they want us to say." The evidence is that during this part of the episode, the students are viewing knowledge as consisting of remembered rules and vocabulary rather than personally constructed or authentic understanding.

There are exceptions to the above characterization of students in the episode. There is a moment in which the students raise ideas, when Luisa asks "Shouldn't it gain ten?" and Hannah agrees, saying "because it's twice as heavy;" Hannah is briefly more animated at this moment, speaking loudly, looking directly at Luisa, and pointing at her with her pen. The moment quickly passes, though, and the students' attention drops back to the worksheet. A moment later, when Yvonne says "Ah, shoot," there are a few seconds during which the students smile, laugh, shrug, and touch their own faces or fix their hair. Again, though, the students' attention returns to the worksheet after a brief period.

At the moment indicated by the solid horizontal line in the transcript, however, the students' behavior changes more definitively: they sit up and make eye contact with one another, speak in relatively clear, loud, dynamic voices, and gesture more prolifically. No students rest their foreheads on their hands, although one rests her chin. At that same moment, the students' conversation changes to original speech about the physics ideas in question, starting with Yvonne's statement of her intuition in contrast to the answer she had heard the group decide ("See intuitively, I would think that it'd…"). Erin and Hannah interject briefly and then Yvonne continues, explaining her reasoning that if the truck slows down the car has to speed up. During what follows, Erin and Hannah both gesture simulations of the collision. Hannah's gesture is part of her reasoning about a situation she poses for comparison, a situation not on the worksheet, a move Luisa and Yvonne evidently consider appropriate (judging by their chiming in at the end of the episode). In sum, during the period of more animated behaviors, the substance of students' conversation is the physical events under consideration. (The episode continues for another 12 conversational turns before the students shift their focus back to the worksheet).

The shift to conversation in which the students work together to identify the salient elements of the physical scenario is precipitated by a conversation overheard from the next table: a TA says "So what would you think intuitively?" and a student begins her response with "Intuitively I was thinking that…." Hannah seems to be responding to that conversation when she says "I hate that word, *intuitively*." We may only speculate about the reasons for her distaste. Perhaps she resents the vulnerability inherent in revealing her personal ideas; alternatively, she may suspect that the worksheet authors consider her intuitions to be both predictable and incorrect. In any case her comment highlights the tension between answering authentically and answering correctly.

The observation that changes in behavior correspond with changes in the substance of the students' verbal exchanges, in this and other episodes, supports our identification of the students' epistemological framing. In this case, students shift from framing their activity as *completing the worksheet* (leaning over the worksheet, reading it aloud, and confirming one another's brief responses) to framing it as *discussing ideas* (animatedly figuring out how Newton's third law might apply to this physical situation). We identify these framings not on the basis of the students' behaviors alone or the substance of the students' discussion alone. Instead, the evidence of framing is in the coherence between the behaviors and the substance.

B. Developing and applying a systematic observational protocol

Collaborative analysis generates hypotheses that ideally send individual researchers back to the data corpus for further investigation. Analyses such as those described above demonstrate that there is evidence of framing in easily ob-

served features of students' behavior. Systematic analysis of more groups and other tutorials led to the discovery of four distinct behavioral patterns. In some segments of video, for example, the students are bent over their worksheets, talking in subdued voices with their hands relatively still; in other segments, they are sitting up straight, looking at each other, speaking in loud voices, and gesturing prolifically. These clusters of co-occurring behaviors stand out clearly; having had them described and with very little training, independent coders achieve 90% agreement in classifying 5-second video segments into one of four clusters. Thus, our collaborative analysis generated a coding scheme.

In what follows we present the method used to create this coding scheme—that is, to identify these behavioral clusters. We then discuss the meaning that these behaviors appear to hold for the participants. Finally, we share preliminary results of comparative and correlational studies inspired by the coding scheme.

1. Identifying patterns

In order to identify clusters of behaviors that appear in our data set, coders independently watch video of small groups of students and note the exact time of *changes in* the students' vocal register, affect, grammar, gesture production, and body language. Having noted changes in behavioral clusters, we then systematically identify the features of the behavioral cluster: hand motions including gestures,[27] facial aspect, body position and/or movement, vocal register, gaze, and so on. We label the behavioral clusters with meaning-neutral labels (colors). Identifying the behaviors within each cluster permits coders to reliably observe when students participate again in a behavioral cluster the coder had previously identified (i.e., shift back into a certain cluster after engaging in other behaviors).

Coding is performed in real time without transcript. At least two researchers code the changes independently. Earlier research led us to expect that clusters of behaviors typically change all at once for an entire group of participants[20] and that is typically the case in our data. Coders are able to code the whole group reliably (peers typically share the same behavioral cluster).[35] The behavioral clusters are distinct and easy to identify, with over 95% inter-rater reliability on the identification of the cluster type and 90% inter-rater reliability (before discussion) on the timing of the transitions, to 5 s accuracy.

Working in this way, we identified four clusters of behaviors that applied to all three groups we studied. The fact that the behavioral clusters identified for one group appeared in others as well is a finding in itself, probably reflecting a set of cultural norms shared by the participants (almost all of whom are junior and senior university students).

(a) *Blue behavioral cluster.* In the behavioral cluster that we have labeled "blue," students' eyes are primarily on their papers with brief glances up at their peers. Their bodies lean forward at an angle of about 30° to the vertical, their hands are mainly at rest (few gestures), and their faces are relatively neutral; the opening frame of Video 1 is typical. Their tone of voice is low, quiet, and indistinct (muttering); when they speak to one another, they briefly and quietly speak with

Video 1 (Color) Verbal and nonverbal cues indicate reading from a shared text.

rising intonation and answer briefly and quietly with falling intonation. The initiator often utters an incomplete sentence that is completed by a peer.

(b) *Green behavioral cluster.* In the "green" behavioral cluster, students sit up straight and make frequent eye contact with one another, as shown in the opening frame of Video 2. Their faces and voices are animated; they gesture relatively prolifically and engage in clear, loud, dynamic speech.

(c) *Red behavioral cluster.* The "red" behavioral cluster involves the students' interaction with a teaching assistant, although students need not exhibit the red behavioral cluster all the time that a TA is present. (Sometimes, for example, they exhibit green behaviors with a TA nearby or even joining the discussion.) The students speak little and make eye contact with the TA. They sit straight with their bodies still, and their hands are relatively quiet except for habitual movements.

(d) *Yellow behavioral cluster.* In the "yellow" behavioral cluster, students giggle or smile and have hedging or joking tones of voice. They shift their bodies around in their seats and touch their own faces and hair. Their gaze is unsettled, moving among peers, papers, and other points in the room. The yellow behavioral cluster appears to be somewhat less stable than the other three clusters identified in our examples: although it recurs regularly, it rarely lasts even as much as thirty seconds, whereas the other three clusters often persist for minutes at a time.

Video 2 (Color) Changes in behavior correspond with changes in the substance of the students' verbal exchanges.

TABLE I. Specific behaviors found in each cluster and expectations likely to be associated with those behaviors.

Blue: Worksheet Frame		Green: Discussion Frame	
Behavior	Expectation	Behavior	Expectation
Hands quiet, face neutral	Minimal interaction, individual activity	Prolific gesturing	Peers are watching and want to understand
Body leans forward, eyes on paper	Attention belongs on the worksheet	Animated tone, face	Intellectual and/or emotional engagement
Brief glances at peers	"Check-ins" expected	Sit up straight, eye contact	Attention belongs on peers
Muttering	Peers not attending to details of speech	Clear utterances	Peer interest in details of speech
Red: TA Frame		Yellow: Joking Frame	
Behavior	Expectation	Behavior	Expectation
Sit up straight; eye contact with TA	Attention belongs on TA	Giggle, smile, self-touch, fidget, unsettled gaze	Embarrassment, perceived vulnerability
Reduced gestures	Rehashing thinking		

A summary of the features of the four behavioral clusters appears in Table I.

2. Interpreting patterns

We labeled the behavioral clusters with colors in an attempt to suspend our inclinations to interpret them while we defined the specific physical behaviors involved. Having arrived at these definitions and shown inter-rater reliability, we shifted to interpretation: what do these different behavioral clusters suggest with respect to epistemological framing? For us, as we expect for the students, this interpretation is automatic and unconscious—it is difficult not to do it. That is reason for attending to it systematically in analyzing student thinking.

(a) *Blue behavioral cluster.* The behaviors associated with the blue and green behavioral clusters indicate very different framings of the tutorial activity in spite of the fact that both might loosely be called "working collaboratively on the tutorial." In the blue behavioral cluster, the fact that students' eyes are on their papers and their bodies lean over the desks sends the message that their main interaction is with the worksheet. Students' neutral facial expressions and lack of gestures further suggest that they expect only a modest level of interaction with their peers. They occasionally glance at one another, indicating an expectation of "check-ins" with peers, and they do often speak to one another, typically reading from the worksheet, giving brief status reports of their progress, and requesting or providing information. The students evidently frame the activity as primarily *completing the tutorial worksheet.*

(b) *Green behavioral cluster.* The green behavioral cluster, on the other hand, is characterized by speech delivered in relatively loud animated voices, indicating intellectual and/or emotional engagement. Sitting up straight and making eye contact with one another, students display an expectation that their attention belongs on their peers. Their clear utterances

and prolific gestures indicate further that they expect their peers to pay attention to what they are trying to express. Watching students behave in this way, we understand them and they understand each other to be engaged in a *discussion.*

(c) *Red behavioral cluster.* In the red cluster, students display expectations that their attention belongs on the TA, suggesting that they frame the activity as *listening to the TA.* Variations in the red behavioral cluster sometimes include students responding to TA requests for explanations of phenomena that the students had already discussed among themselves. In those cases, the red behavioral cluster includes students speaking more dynamically and gesturing. The verbalizations and gestures, however, have a different quality than those observed in the green frame; the gestures, for example, are often reduced versions of gestures originally made during the green behavioral cluster.[30] Students' attention remains primarily on the TA, as indicated by gaze and body orientation. In those cases, however, it might be more accurate to label the activity as *responding to* (rather than listening to) *the TA.*

(d) *Yellow behavioral cluster.* The yellow cluster of behaviors is less straightforward to interpret than the others. On the one hand, they seem to indicate social discomfort or perceived vulnerability: students avoid direct eye contact, fidget in their seats, and fiddle with objects. On the other hand, they laugh, smile, and appear to joke with one another, behavior that indicates play. These characterizations are not mutually exclusive; play is a "nonliteral orientation"[20] that may be initiated as a cover for actions involving some personal risk. This behavioral ambiguity can correspond to ambiguity in the communicated substance of student thinking.[1] We refer to the frame indicated by the yellow cluster as the *joking frame* to include the sense that students are signaling that the content of their speech should not be taken literally (even though they may also intend, on some level, to explore the ideas they express).

Table I details the behaviors characterizing the different behavioral clusters and the expectations we take them to indicate. To be clear, we are describing our interpretations as what we expect is shared cultural knowledge, shared among students, researchers, and most readers of this journal. A fuller analysis supports these interpretations.[1]

3. Comparing and correlating patterns

Using data from three groups of four students, we found four distinct behavioral clusters, all of which involve communication among the participants. They arose from student behavior, not from top-down categorization, and account for nearly all of the time that students spend engaged in the collaborative active-learning activities under study. The same behavioral clusters were identified for all four groups, suggesting that the four frames indicated by the behavioral clusters are common to many tutorial experiences.

(a) *Comparison of patterns of behavioral clusters for tutorial groups in class periods.* The behavioral coding suggested simple comparisons across groups and tutorials. For example, we coded Jasmin's group's work on the entire Newton's third law tutorial, including the short segment described above, noting the temporal sequence of behavioral clusters and the fractions of cumulative time for each cluster during small-group work. We also coded Hannah's group working on the tutorial on Newton's third law (including the episode described above), Hannah's group working on a tutorial on free-body diagrams, and a third group doing a tutorial on electrostatics. These different groups and different activities show different patterns of behavioral clusters in the course of a class. Jasmin's group, for example, spends about 50% of its time in the green behavioral cluster during the Newton's third law tutorial and about 30% in the blue; the pattern is similar for the group doing the electrostatics tutorial. However, Hannah's group, doing the third law tutorial, spends only 20% of its time in the green behavioral cluster and about 65% of its time in the blue. Hannah's group shows a different pattern of behavioral clusters in the tutorial on free-body diagrams, spending about a third of the time in each of the blue and green behavioral clusters. The yellow behavioral cluster accounts for only a small fraction of each group's time or, in one case, does not appear at all.

(b) *Correlation between green behavioral cluster and mechanistic reasoning.* In addition to closely analyzing the interaction between student behavior and the substance of conversation, as illustrated above, we have examined some of the data for statistical correlations. Specifically, we have looked for correlations between the green behavioral cluster (discussion frame) and reasoning about the causal mechanisms that underlie natural phenomena. As defined by previous research,[36] mechanistic reasoning about a physical phenomenon involves several elements: describing the target phenomenon, identifying setup conditions, identifying entities, identifying actions, identifying properties of entities, identifying the organization of entities, and "chaining" (verbally expressed logical inference). The coding scheme based on these elements provides a systematic means of analyzing student conversation for evidence that they are reasoning about physical mechanism. In that scheme, the most convincing evidence of mechanistic reasoning is chaining. For example, in discussing Newton's third law with Jasmin, her partner Alan says "Cause if there's equal and opposite forces, the truck… we would, if we actually saw it, we'd think the truck would hit the car and go backwards because of the force." Alan's statement contains evidence that he is reasoning about how the collision works: specifically in the categories of entities (the truck and the car), activities (the truck would hit the car, the truck would go backward), and chaining, in the student's inference that if the forces were equal and opposite then the truck would go backward as a result of the collision—and therefore the forces cannot be equal and opposite.[37]

In a pilot project,[38] 20-minute video episodes from six tutorial sessions chosen at random were coded for behavioral clusters (without transcript, as described above). In a separate analysis, the sessions were transcribed and coded for (a) instances of some level of mechanistic reasoning and (b) instances of chaining. Two independent coders agreed on 90% of the behavior codes and 87% of the mechanistic reasoning and chaining codes before discussion. The codes for behavior, mechanistic reasoning, and chaining were then matched for each 5 seconds of the tutorial session. Looking across groups, 81% of the chaining occurred during the discussion frame (green behavioral cluster). This result supports the analyses above that suggest interaction among the students' behavior, the substance of their reasoning, and their framing of what is taking place.

Our conjecture is that this correlation reflects a consonance between students' framing their activity as discussion and their reasoning about causal mechanisms, a consonance that may occur for several reasons. First, gesturing can help individuals reason about mechanisms by performing simulations or depictions of the entities and their organization and activity; second and similarly, they provide a means for communicating those ideas to others.[30] For these reasons, a frame that affords gesturing may be more conducive to mechanistic reasoning, and students interested in discussing mechanism may be more inclined to such a framing. Another possibility is that the correlations we find between the discussion frame and mechanistic reasoning reflect more simply the fact that to engage in chaining requires more extended conversational turns and more extended attention by conversational partners. In either case, behaviors in the green cluster are associated with scientifically valuable reasoning and interactions.

V. IMPORTANCE OF BALANCING CODING WITH INSIGHT

It can be tempting to think that the goal of video analysis is to develop a coding scheme. Collaborative insight-oriented analysis of short episodes, after all, deals with only a very small fraction of the total time that students spend in the classroom, and the analysis process sometimes emphasizes interpretational diversity over uniformity. It is primarily by engaging with a significant fraction of the data corpus that we address questions about the distribution and frequency of events observed in specific episodes. Coding schemes are how we scale up our detailed analyses and establish reliability for particular measures.

Much is lost, however, in the change of scale. Coding schemes by their nature reduce rich interactional sequences to one or two component parts and suggest possibly spurious correlations or even causal relations between those parts. In the color-coding analysis above, for example, I am tempted to think of Jasmin's group as "better" because they spend more time in the discussion frame than Hannah's group. But what does it mean that they spend their time that way? Does "time on task" even indicate the importance of a particular framing to the tutorial experience? Does it contribute to their success in tutorials (however that might be defined)? We would need to return to the video to carefully establish our confidence in such interpretive moves.[36] In another example, earlier work by our research group established that the majority of student-TA interactions in tutorial are initiated by teaching assistants.[23] However, the analysis revealed nothing about the meaning of that result. Is the quality of the interaction related in any way to who initiates it? An analysis that categorizes thousands of events is only as meaningful as each single event of which the data are constituted.

[1] R. Scherr and D. Hammer, Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics, Cogn. Instruct. **27**, 147 (2009).

[2] B. Jordan and A. Henderson, Interaction analysis: Foundations and practice, J. Learn. Sci. **4**, 39 (1995).

[3] J. Johnson, Mixing humans and nonhumans together: The sociology of a door-closer, Soc. Probl. **35**, 298 (1988).

[4] B. Latour, *Reassembling the Social: An Introduction to Actor-Network Theory* (Oxford University Press, New York, 2005).

[5] S. Sismondo, *An Introduction to Science and Technology Studies* (Blackwell, Malden, MA, 2006).

[6] C. Goodwin, Action and embodiment within situated human interaction, J. Pragmat. **32**, 1489 (2000).

[7] E. Hutchins, *Cognition in the Wild* (MIT Press, Cambridge, MA, 1995).

[8] E. Hutchins, How a cockpit remembers its speeds, Cogn. Sci. **19**, 265 (1995).

[9] *Communication and Cognition at Work*, edited by Y. Engeström and D. Middleton (Cambridge University Press, New York, 1993).

[10] E. F. Redish and D. Hammer, Reinventing college physics for biologists: Explicating an epistemological curriculum, Am. J. Phys. **77**, 629 (2009).

[11] L. C. McDermott and P. S. Shaffer, *PEG, Tutorials in Introductory Physics* (Prentice-Hall, Upper Saddle River, NJ, 1998).

[12] L. C. McDermott, P. S. Shaffer, and M. Somers, Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood's machine, Am. J. Phys. **62**, 46 (1994).

[13] R. M. Goertzen, R. E. Scherr, and A. Elby, Accounting for tutorial teaching assistants' buy-in to reform instruction, Phys. Rev. ST Phys. Educ. Res. (to be published).

[14] R. M. Goertzen, R. E. Scherr, and A. Elby, Tutorial TAs in the classroom: Framing teaching as a focus on indicators, (unpublished).

[15] R. M. Goertzen, R. E. Scherr, and A. Elby, *Indicators of Understanding: What TAs Listen for in Student Responses*, Physics Education Research Conference, AIP Conference Proc. No. 1064, edited by C. Henderson, M. Sabella, and L. Hsu (AIP, Edmonton, Alberta, 2008), pp. 119–122.

[16] A. Elby, Helping physics students learn about learning, Am. J. Phys. **69**, S54 (2001).

[17] A. Elby, R. E. Scherr, T. McCaskey, R. Hodges, E. F. Redish, D. Hammer, and T. Bing, *Open source tutorials in physics sensemaking*, 2007.

[18] G. Bateson, in *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*, edited by G. Bateson (Ballantine, New York, 1972), pp. 177–193.

[19] E. Goffman, *Frame Analysis: An Essay on the Organization of Experience* (Northeastern University Press, Boston, 1986).

[20] D. Tannen, *Framing in Discourse* (Oxford University Press, New York, 1993).

[21] G. L. MacLachlan and I. Reid, *Framing and Interpretation* (Carlton, Victoria, 1994).

[22] J. Roschelle, B. Jordan, J. G. Greeno, B. Katzenberg, and C. Del Carlo, *Preliminary Report on Classroom Observations for the National Board for Teacher Certification* (Institute for Research on Learning, Palo Alto, CA, 1991).

[23] R. E. Scherr, R. S. Russ, T. Bing, and R. Hodges, Initiation of student-TA interactions in tutorials, Phys. Rev. ST Phys. Educ. Res. **2**, 020108 (2006).

[24] Frames, scripts, and schemata are related and overlapping terms in the fields of linguistics, artificial intelligence, cognitive psychology, social psychology, sociology, anthropology, and other disciplines; an overview and history of the uses of these related terms appears in Chap. 1 in Ref. [20].

[25] Linguistic markers include omissions, repetitions, negatives, modals, etc., and are described in Chap. 1 in Ref. [20].

[26] There is the possibility of the speaker acting as though she is engaged in a lively discussion while in fact not experiencing genuine engagement. For frame analysis in the context of acting, see Ref. [19].

[27] We have developed a variation in this pattern for meetings in which one or more members are participating from a remote location by videoconference: each participant obtains the video episode in advance. During the meeting, group members "sign off" to watch predetermined short segments of the episode independently at their different locations (usually a minute at a time) and then reconvene for discussion of that segment.

[28] A subtle effect of episode selection is that the length of an episode can affect the time scale considered significant for analysis. In an episode a few seconds long, the contribution of subsecond interactions is sometimes more visible than over the course of an hour.

[29] Students' names are pseudonyms.

[30] R. Scherr, Gesture analysis for physics education researchers, Phys. Rev. ST Phys. Educ. Res. **4**, 010101 (2008).

[31] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, Phys. Teach. **30**, 141 (1992).

[32] D. P. Maloney, Rule-governed approaches to physics: Newton's third law, Phys. Educ. **19**, 37 (1984).

[33] R. K. Boyle and D. P. Maloney, Effect of written text on usage of Newton's third law, J. Res. Sci. Teach. **28**, 123 (1991).

[34] S. Goldin-Meadow, *Hearing Gesture: How our Hands Help us Think* (Harvard University Press, Cambridge, 2003).

[35] This result is partially to be expected: since peers' activity together is mutually constructed, it is natural that they should share their framing of a given situation and therefore participate in the same behavioral cluster. On the other hand, it seems possible that individuals could "opt out" of the framing mutually constructed by the others in the group. This paper reports brief occurrences of such opting out or mismatch of framings. However, the great majority of the time, peers share the same behavioral cluster.

[36] R. S. Russ, R. E. Scherr, D. Hammer, and J. Mikeska, Recognizing mechanistic reasoning in scientific inquiry: A framework for discourse analysis developed from philosophy of science, Sci. Educ. **92**, 499 (2008).

[37] His reasoning is not correct; see Ref. 36 for a discussion of correctness and mechanistic reasoning.

[38] L. D. Conlin, A. Gupta, R. E. Scherr, and D. Hammer, *The dynamics of students' behaviors and reasoning during collaborative physics tutorial sessions*, Physics Education Research Conference, AIP Conference Proc. No. 951, edited by P. R. L. Heron, L. McCullough, and J. Marx, (AIP, Greensboro, NC, 2007), pp. 69–72.