San Jose State University

SJSU ScholarWorks

Variation of instructor-student interactions in an introductory interactive physics course

Emily West
University of California, Davis

Cassandra Paul
San Jose State University, cassandra.paul@sjsu.edu

David Webb
University of California, Davis

Wendell Potter
University of California, Davis

Recommended Citation

Variation of instructor-student interactions in an introductory interactive physics course

Emily A. West,[2] Cassandra A. Paul,[1] David Webb,[2] and Wendell H. Potter[2]

[1]*Department of Physics and Astronomy, San José State University, One Washington Square, San José, California 95192, USA*
[2]*Department of Physics, University of California at Davis, One Shields Avenue, Davis, California 95616, USA*

The physics instruction at UC Davis for life science majors takes place in a long-standing reformed large-enrollment physics course in which the discussion or laboratory instructors (primarily graduate student teaching assistants) implement the interactive-engagement (IE) elements of the course. Because so many different instructors participate in disseminating the IE course elements, we find it essential to the instructors' professional development to observe and document the student-instructor interactions within the classroom. Out of this effort, we have developed a computerized real-time instructor observation tool (RIOT) to take data of student-instructor interactions. We use the RIOT to observe 29 different instructors for 5 hours each over the course of one quarter, for a total of about 150 hours of class time, finding that the range of instructor behaviors is more extreme than previously assumed. In this paper, we introduce RIOT and describe how the variation present across 29 different instructors can provide students in the same course with significantly different course experiences.

I. INTRODUCTION AND BACKGROUND

Physics education researchers have shown in numerous studies that interactive-engagement (IE) courses promote student learning of physics concepts better than their traditional counterparts [1–3]. These courses are defined as those in which students engage in activities where they make sense of material while engaging in discussion with peers and/or instructors. Because IE courses are for the most part successful at improving student learning outcomes, many physics departments nationwide are adopting IE courses [4–6]. However, there is a large variety in the success of interactive-engagement courses [1]. For some IE courses the student conceptual understanding gains, as measured by the Force Concept Inventory (FCI) [7], are indistinguishable from those in traditional courses. Additionally, while traditional courses typically produce normalized FCI gains tightly clustered around an average of 0.23, an interactive-engagement course can produce gains anywhere from 0.23 to 0.70 [1].

Some of the variation in IE course success is due to the broad curricular definition of the term interactive engagement, but a portion of this variation is also likely due to the pedagogical implementation at the instructor (faculty or graduate student) level [8,9]. That is to say that some of the variation is due to differences in course structure (time devoted to different types of content, balance between lecture and laboratory and discussion,

etc.) whereas other aspects of the variation are due to the manner in which a particular instructor interacts with the class.

It is typical at the university level for many instructors to teach different sections of the same course. This is particularly true for graduate student teaching assistants (TAs). As more IE courses are being implemented [3,10,11], graduate students are becoming increasingly responsible for using interactive pedagogy, because often the IE elements are present in the laboratories and discussion courses.

As of yet, there has been little research conducted with regard to graduate teaching assistant pedagogy. That which does exist suggests that there is a variation in types of interactions the instructors engage in. A former member of the UC Davis physics education research group completed dissertation research on TA actions in the classroom, defining several different typical instructor modes ranging from observer to obtrusive [12]. Additionally, Goertzen *et al.* found that not all TAs have the same amount of buy-in to course philosophy, which can affect how they implement the course material [13]. They found evidence that TAs use different indicators to determine if their students are making sense of the material [14]. Finally, emerging research shows that these differences, specifically instructors' responses to student questions, can stifle or promote student reasoning [15]. These findings suggest that students can have vastly different experiences in different sections of the same course taught by different instructors.

There are ways to measure the variation of instructor implementation, for example, the reformed teaching observation protocol (RTOP) was developed to measure

the extent to which interactive and student-centered techniques are utilized in a given classroom [16]. The RTOP is a 25 item instrument that allows an observer to rate the extent to which particular IE elements are present in a lesson. The RTOP is an excellent tool for evaluating reformed classrooms; however, the RTOP assumes that the instructor is also the curriculum or lesson designer, which is not always the case. In higher education, often graduate student teaching assistants are teaching the IE component (as in the popular University of Washington tutorials, for example). Furthermore, the overall RTOP score does not give us an illustrative view of the classroom, and while field notes may accomplish this to some extent, it is limited to what the observer decides to write down.

We currently lack both a feel for the pedagogical variation that exists in IE courses and a tool that illustrates the difference. Our research fills an important gap in the literature by describing the teaching practices of 29 instructors in a long-standing reformed course supplemented by an extensive TA training and professional development experience for the graduate student instructors. In this paper we examine the pedagogical landscapes that exist in an IE course where all instructors guide the students through identical activities, with identical sets of instructor notes. We use the real-time instructor observing tool (RIOT), a computer program that allows us to code during actual classroom observations, to time and quantify the instructor-student interactions in the classrooms of the 25 TAs and four faculty members teaching the long-standing introductory IE course at UC Davis.

Defining the pedagogical variation among instructors in a course where the curriculum and student population are constant is the first step in determining if this difference has an impact on student learning outcomes.

II. SAMPLE AND ENVIRONMENT

Observations take place in a large-enrollment three-quarter introductory physics course for nonmajors at UC Davis with more than 1500 students enrolled each academic year. The course [5,6,17,18] was designed and institutionalized in the mid 1990s in coordination with educational research findings, and has previously been presented and published simply by referencing the course name, Physics 7. The course design has recently been renamed as Collaborative Learning through Active Sense-making in Physics (CLASP) to more meaningfully capture the course approach and goals. The class is nominally calculus based, though calculus is used sparingly throughout the course. All students at UC Davis taking this level of physics enroll in a CLASP course; there is no traditional alternative, so these students are not a selective group. In a CLASP course, students are expected to use model-based reasoning [19–22] to make

sense of physical phenomena. This sense making is done in an interactive "discussion or laboratory" (DL) which is an environment where 25–30 students work in small groups of five at tables and blackboards. While the inclusion of model-based reasoning activities is one of the elements of the course that is believed to make it effective [5,6,18], this paper will concentrate on the interactive-engagement aspect of the course. Students in the course spend 5 hours per week in bi-weekly discussion or laboratories, while only attending lecture 80 minutes a week (25 minutes of which are typically reserved for weekly quizzes).

The CLASP discussion or laboratories are where the students are expected to learn the bulk of the course content. Learning takes place through a series of activity cycles in which student group work is spliced with instructor led whole class discussions (WCDs). During the small group (SG) time, students are expected to work together to respond to activity prompts on their group's blackboard. The instructor monitors progress of the whole class simultaneously by scanning the blackboards and can then identify students having trouble. When assisting, instructors are not taught to provide immediate answers, but to help guide students back into a productive thinking space until the students can construct a logical argument explaining the physics of the current activity. The instructor is encouraged to conduct a whole class discussion after some groups have finished and all groups have made significant progress toward a complete argument. During whole class discussions, the instructor and students work together to discuss the major points of the activity and larger implications of the phenomena. At this time instructors may call on individuals or groups to present their work, ask follow-up questions to test for understanding, guide the class to a synthesis of the larger implications of the model, or any combination of these approaches. Ideally the instructor facilitates the discussion, but the students voice the majority of the ideas. In cases of class-wide confusion, students should feel comfortable asking questions and participating actively in the discussion. In both small group time and whole class time, the instructors play the role of guiding the students towards correct use of the models. The above description captures the philosophy of those who created the CLASP curriculum; this study shows the extent to which the philosophy is implemented in practice.

The instructors who teach CLASP DL sections are mostly graduate students, though each quarter a few temporary and/or permanent faculty also lead DL sections. While no instructor who strongly dislikes the CLASP style is forced to teach CLASP courses, approximately half (30 out of 63) of the graduate student teaching positions available are for CLASP courses. All new graduate student TAs (which includes the vast majority of all first year graduate students) teach one of

the three CLASP courses on arrival. Of the 30 instructors in our study, 26 are graduate students. Thirteen of the graduate TAs are in their first year and 13 are in their second year or beyond. Two of the 26 graduate students are associated with the physics education research group. All graduate students serving as DL instructors experience a two-part department-run professional development course. Before teaching, instructors participate in an intense two-day training course. Then throughout their first quarter, professional development continues 1 hour per week, including discussions of pedagogy and several opportunities for new TAs to observe one another and experienced DL instructors. The professional development is aimed at communicating the course philosophy to the graduate DL instructors. The details of these programs are not important to the current discussion, though understanding that all instructors have experienced considerable training and exposure to the course philosophy is important. Four of the DL instructors who taught CLASP discussion or laboratories in winter 2008 hold a Ph.D. and are permanent or temporary lecturers or faculty. These instructors are not required to attend any training or professional development associated with CLASP, but two of these individuals are actively involved in the physics education research group at UC Davis.

III. METHODS

Our study focused on observing the interactive elements of CLASP, so we focused only on the instructors who taught the discussion or laboratories and did not include observations of the 80 minute/week lectures. Each of the 30 instructors in our study was observed twice during the quarter. Every observation lasted the entirety of the DL (approximately 140 minutes). The researcher did not interact with the instructor or the students during the observation, and no feedback was given to the instructors about their teaching practice until the quarter was over and all forms of data collection were completed.

The primary data were gathered using the computerized tool RIOT. At the most basic level, the observation protocol differentiates between four major types of interactions: talking at students, shared instructor-student dialogue, observing students, or not interacting. These original types of interactions were chosen years ago to capture TA behaviors of interest to the CLASP course designers. Some form of the RIOT has been used during professional development for years at UC Davis. The original observation protocol required a once per minute observation of the instructors' action at that instant, ignoring the other 59 seconds. Feeling that this eliminated much useful data, we retooled the protocol to use real-time observations through a series of scripts within File Maker Pro database software. During data collection, observers operate on a self-imposed

10 second delay. That is, when an instructor switches from one type of interaction to another, the observer waits 10 seconds before recording the change to make certain that the type of interaction is understood. Any observation taking less than 10 seconds is ignored on the principle that it will not have a lasting impact on the resulting classroom culture. Before beginning data collection, we also subdivided the major interaction types into smaller categories to more richly describe the types of interactions within the classroom. These categories were chosen during preliminary observations and emerged from the data we collected. A summary of all categories of interactions is shown in Table I. The instructor can perform each category of interaction with an individual student, one of five or six different small groups, or with the whole class. Each different combination (category and with whom the interactions take place) appears in a two-dimensional color-coded grid of buttons on the laptop screen. By design, the instructor can only engage in one interaction category at a time. A screen shot of the program interface is included in the appendix (Fig. 6).

All observations conducted during the study were coded using RIOT by one of two researchers. Prior to the start of the study, the two researchers observed the same classroom on three occasions in order to establish inter-rater reliability. Between each observation, the two collaborated extensively to establish a reliable definition of what each category included and excluded. In order to ensure consistent coding, the observers also coded the same classroom for approximately 45 minutes every week. They continuously referred to the category definitions and updated them as necessary. For the length of time corresponding to an entire DL period, 140 minutes, inter-rater reliability varied between 85% and 95%.

The main CLASP course series is three quarters long with parts A, B, and C. During the quarter of observation, five independent CLASP courses were taught: two A and two B courses, but only one C course. In each of the 300-student courses there are 11 discussion or laboratory sections, typically with five TAs teaching 10 of the 11 DL sections and a faculty member or lecturer teaching one section. In each of the five sections (A-1, A-2, B-1, B-2, or C) every instructor was observed during the same DL meeting, enabling direct comparison to all other observations that covered the exact same activities, with two exceptions. See Table II. In addition, it was possible to observe the same DL across the entire sample of all 12 CLASP A instructors (six from each session). This was also true for CLASP B observations, though not for CLASP C since only one session of C occurred during the quarter of the study. Observing all instances of the same DL meeting (across both sessions of A and across both of B) was exhausting to accomplish because they occurred during the same week, so it was only done for one of the two sets of observations: DL 11 for A and DL 13 for B (see Table II).

TABLE I.   A list of all possible instructor-student interactions captured by RIOT.

Type of interaction	Category of interaction	Description	Examples and/or markers
Talking at students	Explaining	Instructor is explaining physics concepts, answers, or processes to student(s)	*Instructor:* "Let me solve problem number 3 for you" *Instructor:* "Heat is a transfer, not a state function."
	Clarifying instructions	Instructor is clarifying the instructions, or reading from the activity sheet, or covering logistical issues, or transitioning between SG and WCD	*Instructor:* "You should have a diagram with five different cases. On the board, write down all five cases and predict which way the electron will deflect." *Instructor:* "The instructions here are a little unclear, what they are asking you to do is figure out the number of modes each of these molecules in each state."
Dialoguing with students	Listening to question	Instructor is listening to a student's question	Student asks a question. Instructor has eye contact with student. Instructor then responds in some way (response coded separately).
	Engaging in closed dialogue	Instructor is controlling conversation, but not explaining. Most commonly, a series of short questions meant to lead the student to the correct answer. Student contribution is one to several words at a time.	*Instructor:* "Was energy lost or gained by the system?" *Student(s):* "Lost." *Instructor:* "So is this an open or closed system?" *Student(s):* "Open" *Instructor:* "What do the energies sum to?" *Student(s):* "Negative Q."
	Engaging in open dialogue[a]	Original definition: A conversation where students contribute about half or more of the words to the dialog. Student may be leading instructor through thinking that had occurred previously or might be actively making sense with the instructor and classmates. Student contribution is complete sentences. Instructor may have initiated conversation with an open-ended question such as "how" or "why." Refined definition: Students are contributing complete sentences, though not actively making sense (sense making moved to ideas, below).	*Instructor:* "So tell me what you've done so far." *Student:* "First we added 50J and 40J and we got 90J. Then we divided by 20sec. We got 4J/s." *Instructor:* "Your units are right, but you should check your calculation." *Student:* "OK, so we divided wrong. We were just rounding."
	Engaging in open dialogue with ideas being shared [a]	A conversation where students contribute about half or more of the words to the dialogue. Student specifically is having a productive conversation with the instructor where concepts are being challenged and worked on. Student contribution is complete sentences, specifically the active development of physics ideas.	*Example 1* A student asks how a wheel can spin if there is no torque on it. The instructor refers students to the linear case, and the students draw the analogy between linear and rotational dynamics in their own words. *Example 2* *Instructor:* "How would you describe what power is?" *Student:* "Hmm, power, it's like a rate, or like the flow of energy. It describes how much flow happens in a given period of time."

TABLE I. (*Continued*)

Type of interaction	Category of interaction	Description	Examples and/or markers
Observing	Passive observing	Instructor is scanning room and assessing classroom progress from afar or browsing board work of groups.	Instructor is not looking directly at individuals, and is not close enough to anyone group to pick out individual conversations. Instructor does not look at any one board for more than 10 seconds at a time. (If they do, this is coded as active observing.)
	Active observing	Instructor is actively listening to individual students or groups.	Instructor is looking directly at students that are talking to each other, or directly at students as they are writing on the board (for more than 10 seconds at a time). Instructor is close enough to hear what individual students are saying and is not reading notes or doing anything else to indicate distracted attention.
	Student presentation	Instructor is listening to students presenting their work to the class.	Students are presenting their work to the whole class, instructor is watching them. (This only happens during WCDs.)
	Students talking serially	Instructor is listening to students talking serially in a whole class discussion. Students are asking each other questions and building on each others ideas. Note: Although we think this is an important event, it happened rarely, and therefore was not considered in the analysis.	After a student makes a presentation, students ask those students questions. OR After the instructor asks the whole class a question, and a student answers, other students jump in one at a time without being called on to contribute to or argue with what the original student has said. (This only happens during WCDs.)
Not interacting	Administrative or grading	Instructor is grading student homework or discussing quizzes, lateness, or other policy.	Instructor is looking at student work at their desk. Instructor may also be telling individual students about their grades or quizzes.
	Working on apparatus	Instructor is helping students with experimental apparatus or computers.	Instructor is setting up the experiment or troubleshooting problems with a computer program.
	Chatting	Instructor is chatting socially with students.	Instructor is engaged in off-topic discussion with students. *Student:* "Do you have a degree in physics? What's grad school like?" *Instructor:* "I have a masters degree, and am working on my PhD."
	Class prep	Instructor is not talking with students or looking at them, but instead getting ready for the next part of the DL.	Instructor is reading notes or writing something on the board while the students are in their small groups
	Out of room	Instructor has left the room.	Self-explanatory

[a]After nearly half the observations were completed (24 of the 30 were observed once), two additional categories were needed. Because this study involves coding in real time, we could not recategorize the data we took previously. One category we added was "Chatting" socially with students. Before adding this category, any chatting was previously categorized under "Administrative or grading." The "Ideas" category was added to split the "Open dialogue" category by distinguishing between students leading the instructor through calculations (unit conversions, for example) and students thoughtfully stringing ideas together with the instructor and classmates. The "Ideas" category is essentially a type of "Open dialogue." To clarify, for 24 of the observations, Ideas is contained in Open, and for 35 observations, Ideas is its own category.

TABLE II. Course and observation breakdown.

Course segment	Lecture session	Number of DL instructors	DL meeting observed
CLASP A	1	6	6, 11
CLASP A	2	6	11, 17, 16 [a]
CLASP B	3	5[b]	4, 13
CLASP B	4	6	3, 13
CLASP C	5	6	7, 13

[a]One instructor was observed during DL 16 instead of DL 11 due to their having a substitute on the planned observation day.
[b]The sixth instructor was dropped from the study because the person teaching the section changed midquarter.

IV. RESULTS

In order to maintain anonymity, each instructor in our study was given a code name corresponding to one of the first 30 elements. Our results reveal a great deal of variation in instructor-student interactions. This variation occurs in the relative amount of small group versus whole class discussion time, the type of interactions that occur during SG time, and the type of interactions that occur during WCD time. Additionally, we find a great deal of independence between the distribution of interactions in SG and WCD time. That is, instructors with similar distributions of interactions in SG time did not necessarily have similar distributions in WCD time, and vice versa (compare Figs. 1 and 2 and Figs. 7 and 8). As a result, we have analyzed the two separately. A larger sample of our data is available in the appendix.

A. WCD versus SG allocation

First, we present the overall time allocation in discussion or laboratory for all 59 observations (see Fig. 3). The average DL section is released 5 minutes early (or started 5 minutes late), taking 135 of the total possible 140 minutes, and there is little variation in the total time spent in DL (standard deviation is 8.8 minutes). On average, the students spend about 40 minutes of the 140 working with the whole class and 95 minutes working in small groups. We find that the student experience in the 140-minute DL sections varies vastly between different observations, with some classes spending as many as
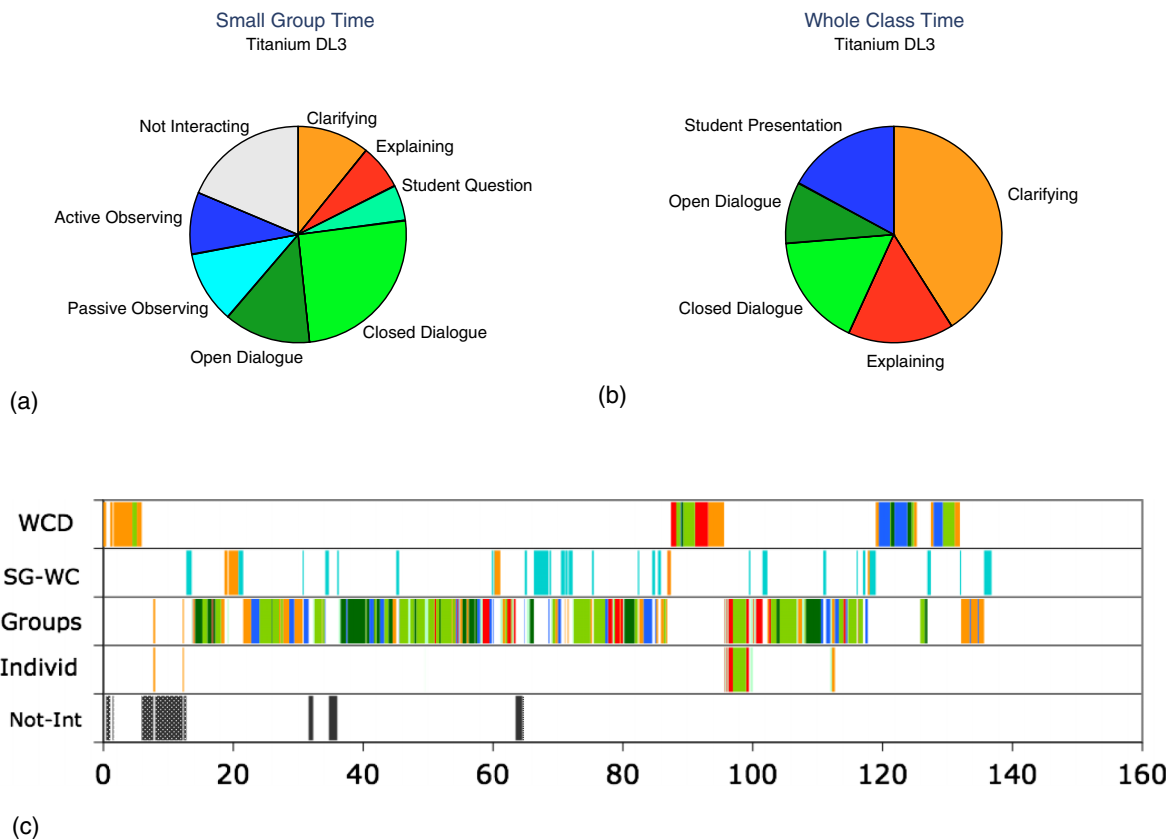

FIG. 1 (color online). Instructor Titanium (DL 3, CLASP B). Small group time breakdown (a), whole class breakdown (b), and the time series data (c). The SG-WC line refers to time when the instructor is addressing the whole class during small group time. By definition, passive observing falls here, but there are other times when the instructor needs to address the whole class briefly without calling for a whole class discussion. Compare with Fig. 2, both instructors are teaching the same set of activities. Notice that they spend their time interacting in similar ways during the small group time, but have a very different set of interactions during whole class time.
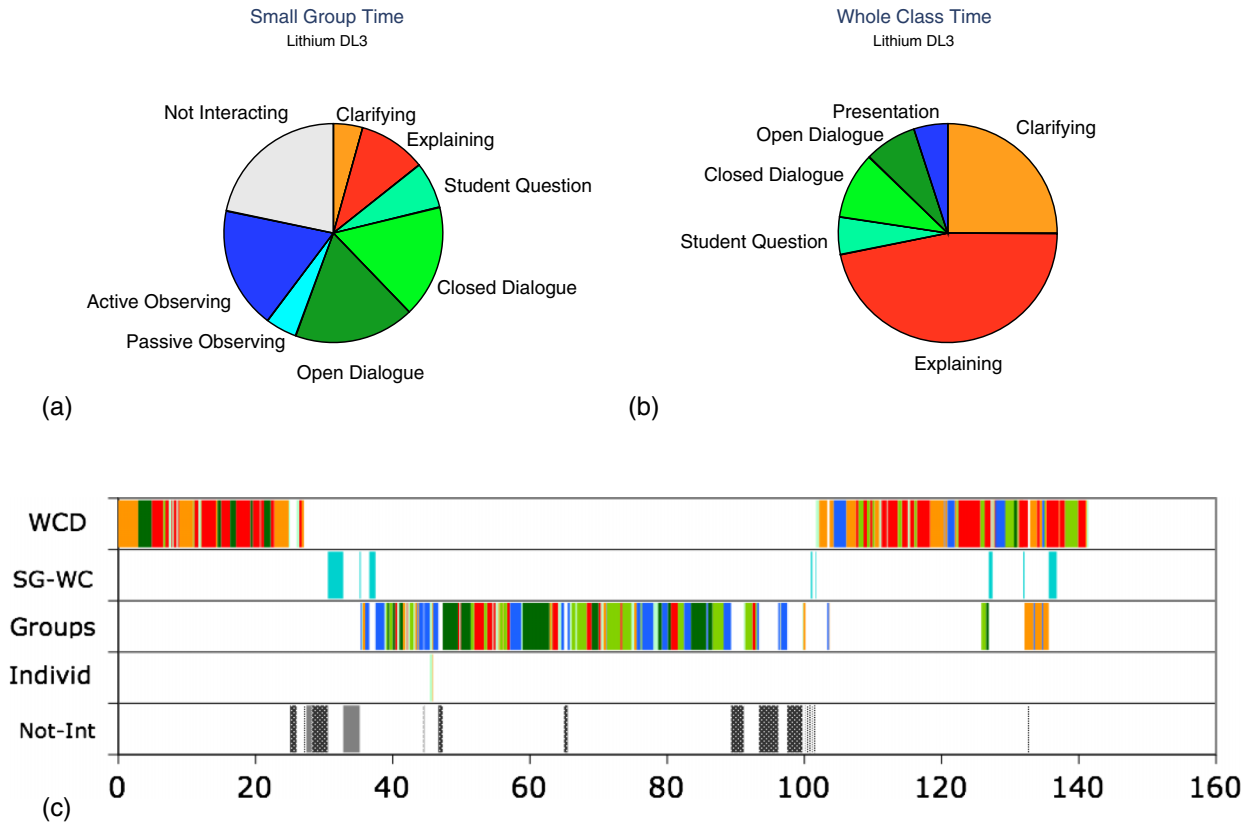
FIG. 2 (color online).   Instructor Lithium DL 3. Small group time breakdown (a), whole class breakdown (b), and the time series data (c). The SG-WC line refers to time when the instructor is addressing the whole class during small group time. By definition, passive observing falls here, but there are other times when the instructor needs to address the whole class briefly without calling for a whole class discussion. Compare with Fig. 1, both instructors are teaching the same set of activities. Notice that they spend their time interacting in similar ways during the small group time, but have a very different set of interactions during whole class time.
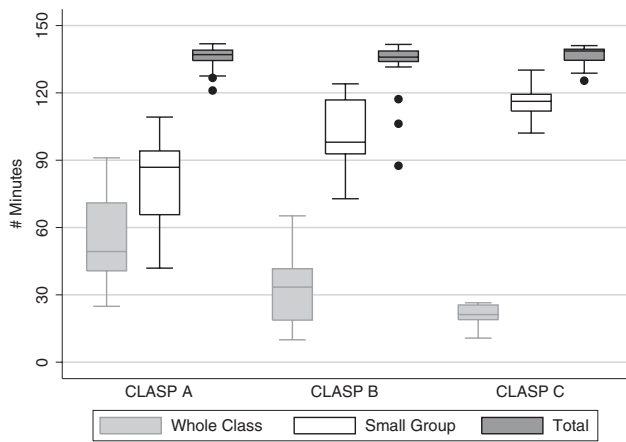

FIG. 3.   Breakdown of time between whole class discussion and small group time is shown by course. Time is in minutes. A complete allotted DL period is 140 minutes. The box contains values from the 25th–75th percentile, with the median marked. The outliers are shown as dots, not connected to the box.

91 minutes working as a whole class and some as few as 10 minutes (standard deviation of 20 minutes). The remainder of the time is spent with students working in small groups, and varies from a minimum of 42 minutes to a maximum of 130 minutes. In other words, some class periods are predominately spent working in a (whole class) group of 25–30 students with the instructor coordinating, whereas others are spent predominately in small groups (4–5 students) with hardly any time as a whole class.

Figure [3] also shows large differences between WCD time allocation in CLASP A, B, and C. CLASP A students, on average, spend considerably more time in WCD mode than the CLASP B or C courses. Indeed, the CLASP A class with the least minutes spent in WCD mode (25 minutes) is similar to the CLASP C class with the most minutes spent in WCD mode (26 minutes), and the 25th percentile for WCD in CLASP A matches the 75th percentile for WCD in CLASP B. The overall amount of time spent in DL is similar across all three courses in the
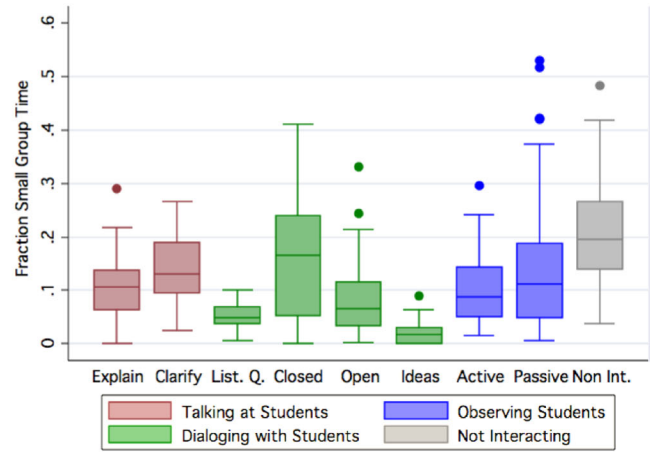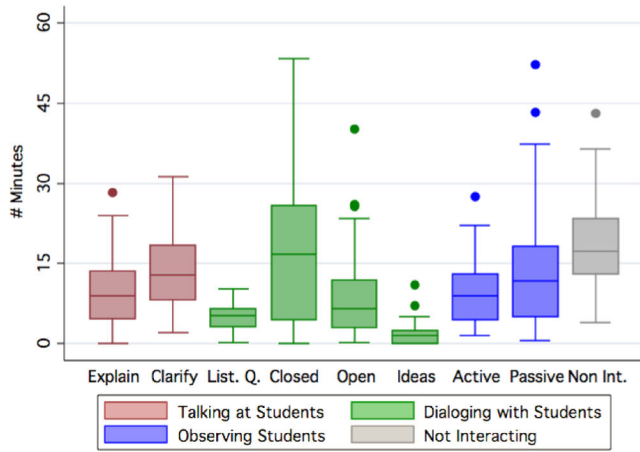
(a)   (b)

FIG. 4 (color online). Breakdown of small group time. (a) The number of minutes of small group time dedicated to each type of interaction. (b) The percentage of small group time in each type of interaction. Note that the category "Ideas" contains only 35 observations because we added the category late. In the other 24 observations all ideas were categorized as open dialogue.

series, though divided differently between small group and whole class time.

B. Small group time

We next turn our attention to the types of instructor-student interactions occurring during small group time. As the overall number of minutes of small group time varies tremendously from observation to observation, results are presented in terms of both the number of minutes instructors spend in certain types of interactions and the percentage of overall small group time instructors spend in certain types of behaviors.

We find that, on average, instructor time is divided roughly evenly between talking at students (24%),

dialoguing with students (31%), observing students (25%), and not interacting with students (20%) (see Fig. 4). However, there is once again large variation across observations. The subcategory of time spent explaining physics content to students ranges from as little as 0 minutes to as high as 28 minutes. Likewise, some instructors spend hardly any time observing students working independently (as low as 4 minutes) whereas others spend as much as 54 minutes. The data are presented in terms of both raw minutes spent engaging in an interaction type and percentage of overall small group time spent engaging in an interaction type. The patterns are similar through either lens, though it should be remembered that the person with the most (or least) raw time spent
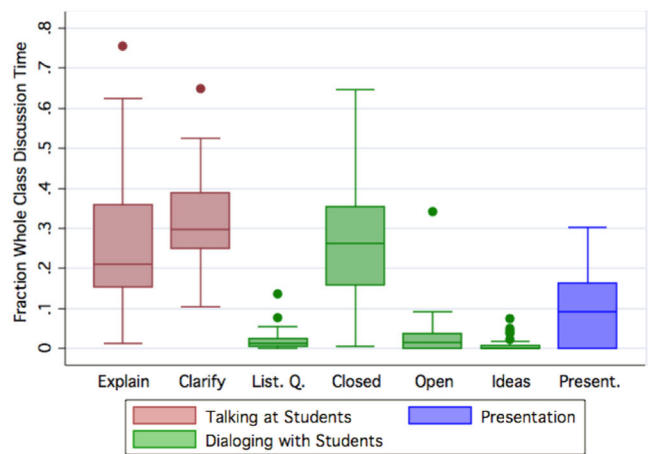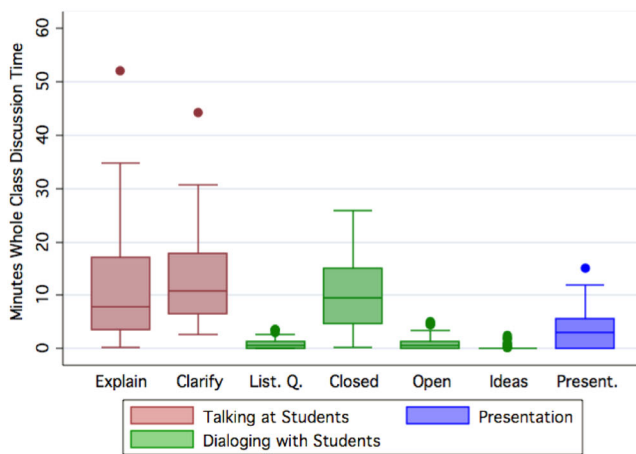
(a)   (b)

FIG. 5 (color online). Breakdown of whole class discussion time. (a) The number of minutes of WCD time dedicated to each type of interaction. (b) The percentage of WCD time in each type of interaction.

participating in a certain interaction is not necessarily the person with the highest (or lowest) percentage of time. The amount of time and proportion of time the instructor spends interacting with students in a way that the students share in the dialogue also has a large range, from as low as 1.2 minutes to as high as 77 minutes, or in terms of percent, from 2.7% to 63%. Not only does the amount of time students spend working in small groups vary, the type of instructor interaction the students experience also spans a large range.

C. Whole class discussion time

The goal of whole class discussion is to provide students with an opportunity to synthesize information and begin solidifying the work they did in their small groups. Like during small group time, instructors use a variety of approaches to facilitate WCD. One technique is to have a student or group of students present the findings from the activity. After the presentation, often the instructor summarizes or rephrases the student presentation or asks other students to comment. Another approach is for the instructor to reframe the questions in the activity sheet in such a way that only a single word or short series of words is required to give the answer, and pose these questions to the class in call and response format. Sometimes the instructor will explain the ideas behind the questions to the students; other times the instructor will retain the short-answer format (which we refer to as closed dialogue) to get the students to fill in key aspects of the overarching ideas. A third approach is for the instructor to convert the discussions into mini lectures, with the instructor predominately explaining to the students. A breakdown of the whole class discussion time by raw minutes and by percentage is shown in Fig. 5.

As with small group time, we see tremendous variation between instructors during WCD. The most common interactions are explaining, clarifying, and dialoguing in closed form. All instructors spend some time clarifying, which is not surprising considering that this category was used to mark transitions between small group time and whole class time, in addition to coding actual clarification. Other than clarification, each of the other modes was never used by some instructors and frequently used by others.

V. DISCUSSION

A. RIOT as a research tool

RIOT was specifically developed using observations from CLASP courses. For this reason, we are not sure the categories we identified will be as useful in other classroom environments. However, RIOT has provided us with invaluable information about what is happening in our classrooms. Researchers interested in using RIOT for their own efforts might first ask themselves if the categories presented here effectively capture what is being looked

for. We recognize that the interactions may look different in other environments and that certain categories might need to be added or removed for different types of classrooms. Our intention is not to showcase an instrument that could eventually be used to measure if certain classrooms meet a ''standard'' of interactivity, but rather to introduce one that illustrates the pedagogical landscape of the classroom for the purposes of internal reflection. This is not to say that RIOT could not eventually be used as a metric for interactivity (in the way RTOP is similarly used), but since we have only observed a single curriculum using RIOT, it has not yet been verified that this is an effective use of the tool.

B. Interactions from the perspective of the curriculum developer

Although the CLASP course philosophy expects certain actions of the TAs, we find a vast range of instructor-student interactions within our interactive-engagement course. The TAs are well exposed to the course philosophy throughout their professional development, and the curricular materials include prompts in the instructor notes meant to reinforce the philosophy, but nonetheless the actual implementation is extremely varied. It is important to note that we did not expect the instructors to interact identically (nor would we want them to) and that, at this point, we are unable to identify which behaviors, if any, are ideal.

For example, excessive ''explaining'' is discouraged in TA training, because students are not necessarily engaged during this interaction, and yet there are some instructors who spend the vast majority of their whole class time doing just that (see B in Fig. 8, Na in Fig. 7, and H in Fig. 8 in the appendix). Additionally, since the students are prompted in the instructor notes to have whole class discussions at specific points in the activities, we did not anticipate that some instructors would omit most of these discussions (see F, Cu, and Li in Fig. 8) while others would spend the majority of time in whole class discussion mode (see Mg, N, and V in Fig. 7). Additional research should be undertaken to see if these deviations from expectation are helpful or harmful to the students.

We do not believe that the instructors purposely ignore their training to teach in an alternative method. Instead, the instructors incorporate the training they receive into their existing framework for teaching, which is typically the result of traditional physics instruction. Each individual TA interprets the curricular goals (as communicated through TA training, twice weekly TA meetings, and instructor notes) in a unique manner, and combines it with their prior experience leading to unique application in the classroom. This observation is consistent with the constructivist teaching philosophy that guides the instruction of students in CLASP, so it is not surprising to find the same type of learning in the instructor population.

At this point, we are not making judgments about what constitutes good or bad teaching interactions. It is possible that the CLASP curriculum is sufficiently robust (i.e., the students have enough opportunities to participate in an interactive environment with their peers) that the individual instructor interactions do not matter. To make such a judgment either we would need to attempt to correlate these interactions with student achievement or we would need to have a better qualitative understanding of the immediate interaction effect on the student. The latter is discussed in the remainder of this section.

C. Rethinking interactions from the student perspective

Because an instructor can only spend limited time with each student or group without ignoring other students, a successful student-instructor interaction should be one where students are able to continue thinking about and making sense of phenomena after the instructor leaves the group. This means that the method used to incite student thinking (which we have captured) is not as important as the resulting student reaction (which we have not captured). It is possible that different instructors use different techniques to procure the same result, and if this result is student sense making, any of those methods are viable. Instead of only coding explain or open dialogue, it would be insightful to code the instructor-student interaction and the resulting student reaction to generate a log of which types of interactions lead to the types of thinking desired. We did not use this approach because it would have required many more observers (there are more students to watch than instructors), and would not have allowed us to observe as many instructors, but this could be the subject of future research.

We have little data on the student actions (student-student interactions) while the instructor is with other groups. When the instructor is observing, we know that all groups in the class are working independently. We should keep in mind that even while the instructor is interacting with a group, about 80% of the class remains working without the instructor. The actions the students take while the instructor is occupied with other groups are certainly important. We suspect that the instructor-student interactions and the resulting classroom culture influence the actions of students while they are not with the instructor, though these data have not been captured. It seems plausible that students who expect explanations when the instructor arrives at their group will behave differently than students who expect to have no interaction with the instructor and differently than students who expect to be prompted by a question from the instructor once the instructor arrives at their group. The observers in this study noted that it seemed like instructors who explained to their students during small group time were more likely to have

students wait for them to come help than to have students who continued trying to work through the more difficult parts of the problem. Unfortunately, because of the nature of the study, no data other than anecdotal were taken in this regard. We have directly observed the instructor interactions, but have not captured data on overall student actions while not in the presence of the instructor; these actions also likely contribute to the classroom culture and student learning opportunities.

VI. CONCLUSIONS

The instructor-student interactions in our physics course vary tremendously. The extent of the variations is surprising given that the graduate student TAs receive two intensive days of training, 10 weeks of continuing TA professional development, and extensive TA notes supporting a particular pedagogy.

The vast interaction differences should be considered when planning curricula that depend on multiple individuals, such as graduate students, for implementation. The instructors responsible for enacting the reformed aspects of the curriculum are not always the same individuals who created the course, so it is important to know to what extent their enactment matches the goals of the course creators. Similarly, it is important for instructors who have enacted curricula developed elsewhere to understand the goals of the curricula used; there is a high likelihood that some percentage of the enactment differs from the vision of the creators.

Instructors are different, and we can depend on them to be so. Graduate student instructors turn over rapidly, teaching for no more than a couple of years. Therefore, even in a long-standing reformed course, with an extensive TA training course, large variations remain in instructor implementation. At this point, we are describing the existence and extent of the variation and not trying to determine why some categories contain more variation than others. The range of instructor interactions ensures that students in different sections of the course experience the course differently. This may not be a problem depending on the resulting student reaction to the interaction. However, we hypothesize that it is possible that explaining (as we have categorized it) fails to promote active sense making from students in the same way that traditional lecture does. Therefore, in order to ensure students in each class have similar learning opportunities, we either need to write these opportunities more explicitly into the curriculum or TAs need more explicit and personalized TA training in their first quarter of teaching.

ACKNOWLEDGMENTS

APPENDIX

In this appendix, we include three figures that provide a little more information about RIOT and the data it collects. In Fig. 6 a screen shot of the program interface gives a sense of RIOT's use. Each category from Table I is combined with who (individual, groups, or whole class) is taking part in the interaction.

Figures 7 and 8 provide additional samples of observation data. Within the paper, we analyze the summarized data, meaning the total amount of time the instructor participates in each type of interaction. Figures 7 and 8 display the data moment by moment. The same patterns summarized in Figs. 3 and 4 in the text can be seen with a careful eye. For instance, the whole class discussion time is the top row of data, and Fig. 7 (observations in CLASP A) shows considerably more time spent in this mode, on average, than Fig. 8 (observations in CLASP B).
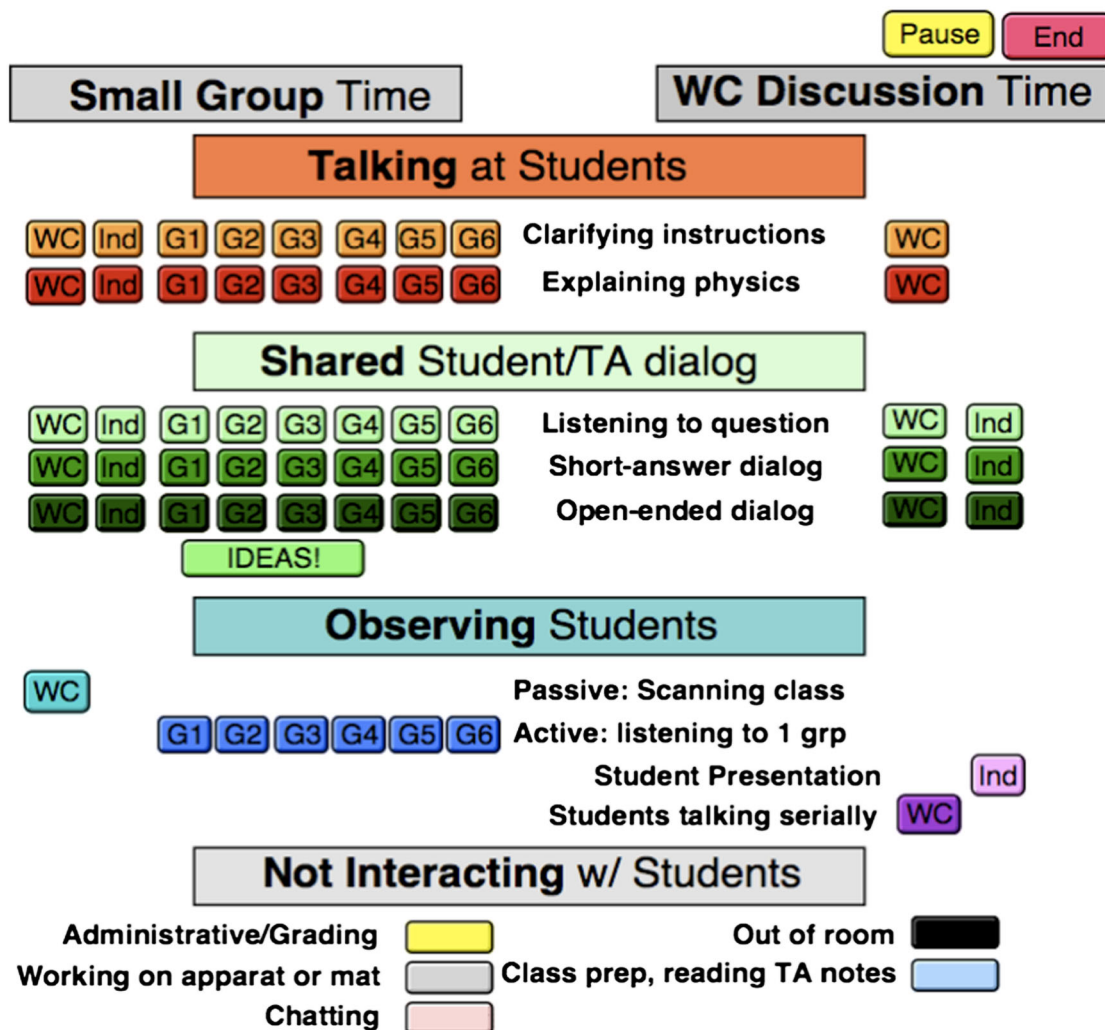

FIG. 6 (color online).   Screen shot of the real-time instructor observing tool (RIOT). The left side of the screen is for when the class is in small group mode, and the right side is for when the class is in whole class mode. The screen is divided horizontally into four larger categories: Talking at Students, Talking with Students, Observing Students, and Not Interacting with Students. Under each of these larger categories are more precise categories. Each row of icons corresponds to the smaller category name that is written next to it. Each small icon labeled WC, Ind, and G1–G6 represents who the instructor is interacting with (whole class, an individual, and groups 1-6, respectively).
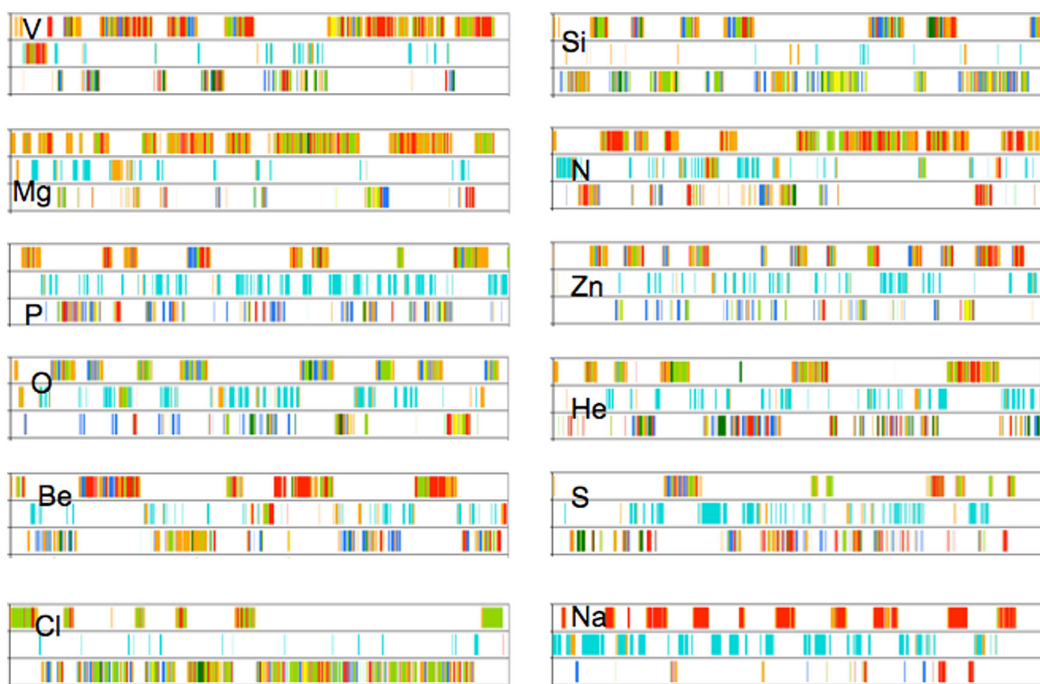
CLASP A Instructors: Observation 1

FIG. 7 (color online).   Twelve different instructors teaching the same set of activities. The top row represents whole class time, the bottom row is small group time, and the middle row represents the interactions that take place during small group time, but when the instructor is addressing the class as a whole. Time is on the *x* axis. The colors are consistent with those in Fig. 1.
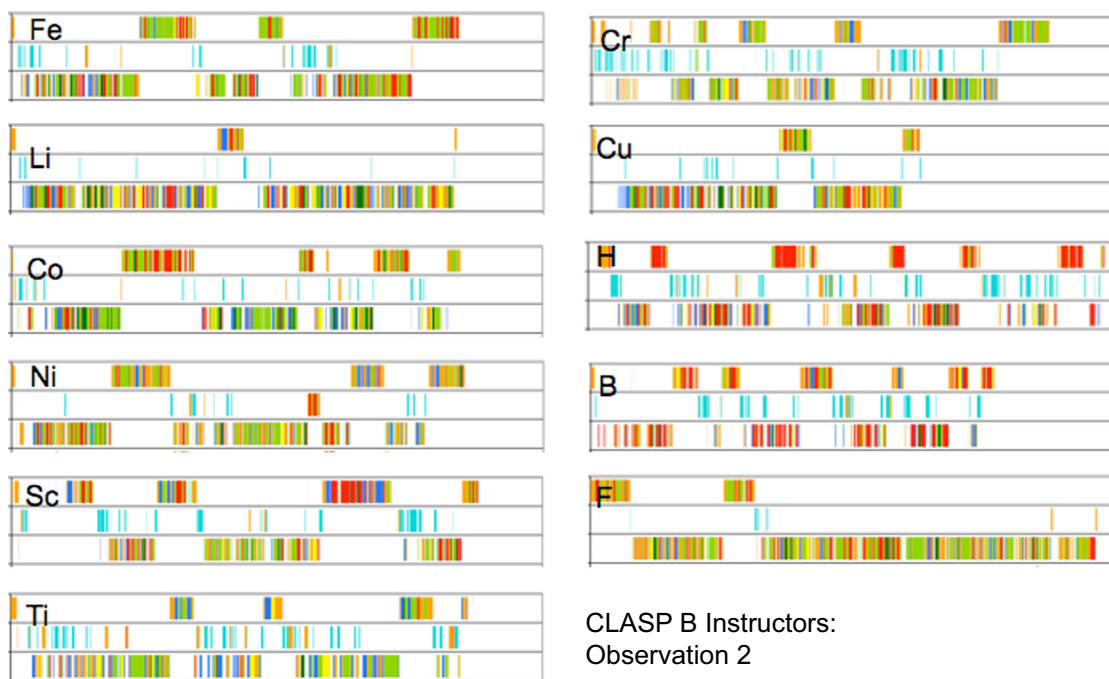

CLASP B Instructors:
Observation 2

FIG. 8 (color online).   Eleven different instructors teaching the same set of activities. The top row represents whole class time, the bottom row is small group time, and the middle row represents the interactions that take place during small group time, but when the instructor is addressing the class as a whole. Time is on the *x* axis. The colors are consistent with those in Fig. 1.

[1] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, Am. J. Phys. **66**, 64 (1998).

[2] M. J. Prince and R. M. Felder, Inductive teaching and learning methods: Definitions, comparisons, and research bases, J. Eng. Educ. **95**, 123 (2006).

[3] K. M. Koenig, R. J. Endorf, and G. A. Braun, Effectiveness of different tutorial recitation teaching methods and its implications for TA training, Phys. Rev. ST Phys. Educ. Res. **3**, 010104 (2007).

[4] L. C. McDermott and P. S. Shaffer, *Tutorials in Introductory Physics* (Prentice-Hall, Upper Saddle River, NJ, 2002).

[5] W. H. Potter, C. DeLeone, and L. Coleman, Radically restructured introductory physics course at a large research university, AIP Conf. Proc. **399**, 829 (1996).

[6] W. H. Potter, D. Webb, E. A. West, C. A. Paul, M. Bowen, and B. Weiss, arXiv:1205.6970.

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

[8] C. A. Turpen and N. D. Finkelstein, Not all interactive engagement is the same: Variations in physics professors' implementation of peer instruction, Phys. Rev. ST Phys. Educ. Res. **5**, 020101 (2009).

[9] C. Paul, E. West, W. Potter, and D. Webb, Graduate-student teaching assistants: A crucial element in improving undergraduate physics, in *APS Meeting Abstracts* (APS, Washington, DC, 2010).

[10] J. D. H. Gaffney, E. Richards, M. B. Kustusch, L. Ding, and R. J. Beichner, Scaling up education reform, J. Coll. Sci. Teach. **37**, 48 (2008) [http://www.ncsu.edu/PER/Articles/JCST_SCALE-UP_article.pdf].

[11] C. Hoellwarth, M. J. Moelter, and R. D. Knight, A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms, Am. J. Phys. **73**, 459 (2005).

[12] A. Calder, Ph.D. thesis, University of California–Davis, 2006.

[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. **5**, 020109 (2009).

[14] R. M. Goertzen, R. E. Scherr, and A. Elby, *Indicators of Understanding: What TAs Listen for in Student Responses* (AIP, New York, 2008), pp. 119–122.

[15] A. Karelina and E. Etkina, Acting like a physicist: Student approach study to experimental design, Phys. Rev. ST Phys. Educ. Res. **3**, 020106 (2007).

[16] D. Sawada, M. D. Piburn, E. Judson, J. Turley, K. Falconer, R. Benford, and I. Bloom, Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol, School Sci. Math. **102**, 245 (2002).

[17] W. H. Potter, C. DeLeone, C. Ishikawa, J. Blickenstaff, and P. Hession, Significant reduction in gender grade disparities in a reformed introductory physics course, in *Proceedings of the 2001 Physics Education Research Conference, Rochester, New York* (American Association of Physics Teachers, College Park, MD, 2001), pp. 111–114 [http://perusersguide.org/items/detail.cfm?ID=4344].

[18] C. DeLeone, W. H. Potter, and G. Potter, Student outcomes in a radically reformed introductory physics course at a large university, in *Proceedings of the April Meeting, National Association for Research in Science Teaching, New Orleans, Louisiana, 2000* (National Association for Research in Science Teaching, Reston, VA, 2000).

[19] D. Hestenes, Toward a modeling theory of physics instruction, Am. J. Phys. **55**, 440 (1987).

[20] I. A. Halloun and D. Hestenes, Modeling instruction in mechanics, Am. J. Phys. **55**, 455 (1987).

[21] R. Frigg and S. Hartmann, Models in science, in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Stanford University, Stanford, CA, 2006).

[22] *Model-Based Reasoning: Science, Technology, Values* (Kluwer, New York, 2002).