Characterizing interactive engagement activities in a flipped introductory physics class

Anna K. Wood,[*] Ross K. Galloway, Robyn Donnelly, and Judy Hardy

Physics Education Research Group, School of Physics and Astronomy, University of Edinburgh,
Edinburgh EH9 3FD, United Kingdom

Interactive engagement activities are increasingly common in undergraduate physics teaching. As research efforts move beyond simply showing that interactive engagement pedagogies work towards developing an understanding of *how* they lead to improved learning outcomes, a detailed analysis of the way in which these activities are used in practice is needed. Our aim in this paper is to present a characterization of the type and duration of interactions, as experienced by students, that took place during two introductory physics courses (1A and 1B) at a university in the United Kingdom. Through this work, a simple framework for analyzing lectures—the framework for interactive learning in lectures (FILL), which focuses on student interactions (with the lecturer, with each other, and with the material) is proposed. The pedagogical approach is based on Peer Instruction (PI) and both courses are taught by the same lecturer. We find lecture activities can be categorized into three types: interactive (25%), vicarious interactive (20%) (involving questions to and from the lecturer), and noninteractive (55%). As expected, the majority of both interactive and vicarious interactive activities took place during PI. However, the way that interactive activities were used during non-PI sections of the lecture varied significantly between the two courses. Differences were also found in the average time spent on lecturer-student interactions (28% for 1A and 12% for 1B), although not on student-student interactions (12% and 12%) or on individual learning (10% and 7%). These results are explored in detail and the implications for future research are discussed.

I. INTRODUCTION

Interactive engagement activities developed through physics education research (PER) have been widely embraced by the physics teaching community [1]. Often used synonomously with the term "active learning," interactive engagement (IE) covers a range of different types of activities from individual problem solving, to working with peers, to interacting with a tutor, and there is now substantial evidence that these teaching approaches lead to better outcomes compared to traditional methods [2,3]. For example, a meta-analysis of 225 studies [3] found student performance on examinations and concept inventories increased under active learning compared to traditional lecturing.

Perhaps the most influential work in this area is a study conducted by Hake involving over 6000 students studying in 62 different introductory Newtonian mechanics courses [2]. Hake measured learning through recording the normalized gain on the Force Concept Inventory (FCI) for each course, and found that those classes which could be described as involving IE methods had substantially higher gains than those in more traditional instruction [2].

However, Hake's results also show that even when courses involve IE, a large FCI gain is not guaranteed. He found that the gains for IE courses ranged from 0.22 to 0.70, whereas the gains for traditional courses ranged from 0.12 to 0.28. This means that for a small number of courses using IE techniques, the gain was actually smaller than the best gain achieved for the traditionally taught courses. This degree of variation implies that the exact implementation of IE can have a large influence on how successful it is. One reason for this may be the way in which instructors implement the pedagogies; for example, Dancy and Henderson found that between a quarter and one-half of instructors deviate significantly from the established design of evidence-based teaching approaches [4]. These results imply that a much more detailed understanding of IE teaching is needed if progress is to be made in optimizing outcomes from these strategies. Research on the efficacy of active learning approaches, such as those described, generally uses a broad definition. For example, Freeman *et al.* [3] describe it as something which

> "engages students in the process of learning through activities and or discussion in class, as opposed to passively listening to an expert. It emphasizes higher-order thinking and often involves group work."

Similarly the definition of "interactive engagement" given by Hake [2] is activities which are

[*]annakwood@physics.org

"designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors…."

While these definitions may be useful for differentiating pedagogies which are essentially reformed in nature from those that take a more traditional, didactic approach, they have been criticized for describing teaching in binary terms [5]. They are also rather broad, encompassing a range of activity types. Further, they give no indication of the frequency with which these activities must be incorporated into lectures, nor for the amount of time that must be spent on these activities for it to count as active learning. Equally problematic is the definition of lecturing, which in its purest form may be thought of as "continuous exposition without interruption" but may also include periods in which the lecturer talks, interspersed with questions to and from the audience. Lecturing is often assumed to be an entirely passive experience, yet these two examples show that the experience of the student can vary considerably. Further, some active learning techniques, such as Peer Instruction (PI), explicitly include time for lecturing [6] as part of the pedagogy.

As research moves beyond showing that active learning has a positive effect, towards unraveling the complexities of how it leads to improved learning outcomes (see, for example, Ref. [7]), a much more nuanced understanding of what actually takes place during active learning pedagogies is needed. Although there has been research on activities in small group physics workshops [8], this has rarely been extended to understanding what takes place in lecture theatre classes. Focusing research on lecture theatre classes is particularly important as this is commonly the primary mode of physics teaching at the university level.

Some research has examined the discourse that takes place in lecture theatre classes [9], but this is predominantly in traditional style lectures and therefore focuses on instructor behavior. There has recently been a call for the interactions that take place in interactive lectures to be studied [10]. This is particularly important as interactions (as the name suggests) are a key aspect of IE strategies. For this reason our approach here, and our major departure from the current literature, is to take a sociocultural approach to analyzing lecture theatre teaching, in which our focus is on the type and duration of students' interactions with the lecturer, with each other, and with the material. IE, therefore, allows us to reconceptualize the lecture theatre as a social space, one in which the student voice is not only heard, but, in collaboration with the voice of the lecturer, encouraged and given value.

Our aim for this work is, therefore, to answer two key research questions:

(1) What types of interactions take place during introductory physics courses utilizing Peer Instruction?
(2) To what extent is each used in a lecture?

We use descriptive statistics to fully characterize the duration and nature of interactions in first year introductory physics lectures at the University of Edinburgh. Through this analysis we propose the *framework for interactive learning in lectures* (FILL), which focuses on characterizing the interactive engagement activities that take place in active learning lectures. The framework is simple enough to be used by researchers with minimal training, but detailed enough to distinguish the key features of IE activities in lectures. It is designed primarily to be used as a research tool that can be used to easily compare different teaching approaches.

II. BACKGROUND

A. Measuring active learning

A number of different ways to characterize classroom activities are reported in the literature. Inventories such as the Teaching Behavior Inventory (TBI) and Reformed Teaching Observation Protocol (RTOP) [11] have been used widely to produce a single measure of the quality of instruction. However, this single quantitative result means that these approaches are less suited to the fine-grained descriptive analysis required for our research. They also do not measure the time or duration of classroom activities.

In response to these problems, two instruments have been developed which aim to provide a more comprehensive and systematic account of activities in the classroom. The Teaching Dimensions and Observations Protocol (TDOP) [12] and the Real-Time Instructor Observation Tool [8] both provide a fine-grained analysis, but with this greater detail comes greater complexity. In order to ensure the validity of such tools extensive training of researchers is required; for TDOP, for example, 28 hours of training is recommended in order to reach an acceptable interrater reliability (IRR).

A simpler system called the Classroom Observation Protocol for Undergraduate STEM (COPUS) [13], designed for use in undergraduate STEM classes, has been developed and aims to overcome this difficulty. Both this system and TDOP aim to document the temporal nature of classroom activities, but they do so by recording codes in 2-min intervals. However, coding in 2-min intervals does not give a precise picture of what happens in the classroom, as many activities, particularly the type of interactions (such as questions) that we are interested in for this research, last less than 2 minutes. This approach therefore limits the usefulness of the instrument.

For these reasons, none of the currently available surveys and instruments for measuring classroom activities are suitable for answering our research questions. In particular, they are not able to provide an accurate characterization of

the type and duration of interactive engagement activities in lecture theatre classes, or to measure changes in activity as and when they happen. Our approach in this research is, therefore, to code interactive activities as they take place on a continuous temporal (second-by-second) basis. This will produce a nuanced and precise quantitative measure of the duration of different types of activity being used during a lecture.

B. Interactive engagement

Activities which are typically counted as involving interactive engagement include individual problem solving, working with peers, and interacting with a tutor or with a computer. These can be introduced into lecture theatre classes through a number of different pedagogies, such as Just in Time Teaching [14], Interactive Lecture Demonstrations [15], and Peer Instruction [16]. Of these, Peer Instruction through the use of electronic voting systems or "clickers" is the best known, and most commonly used by physics faculty [4], and is the technique used in the lectures in this study.

However, the large variation in the type and nature of activities involved in IE pedagogies makes characterization and, therefore, meaningful comparison difficult. For this reason it is common for the shorthand binary terms of "active" and "passive," "constructivist," and "traditional" to be used to describe a teaching approach. Yet as Chi [17] discusses, while the terms active and constructive are used extensively in the literature, only constructive has been explicitly defined. The term active is particularly problematic as it is used to mean many different things. While it is generally a term used to describe activities that involve higher order learning, Chi presents a framework for learning activities in which the active category includes any overt activity such as looking, gazing, gesturing, and pointing. It could also be argued that a student may be actively engaged cognitively in an activity (such as a lecture) that is typically defined as passive.

Despite the broad range of activities that constitute interactive engagement, we believe that the common theme that links them is the idea that these are overt activities which are expected to result in students applying ideas, or engaging in thinking in a way that goes beyond the material that has been provided to them. Our aim here is to take a sociocultural approach based on activity theory (discussed below) to characterize the nature of these activities, and the way in which they are experienced during a lecture. This results in conceptualizing these activities as "interactions." These interactions may occur between students, e.g., peer discussion, between student and teacher, i.e., through asking and answering questions, or through the student interacting directly, and individually, with the material (through problem solving for example). We contrast these "interactive" activities with activities in which students may assimilate information but are not necessarily expected to use this to create

new ideas (such as listening to continuous exposition by a lecturer), which we call "noninteractive." Although it is of course possible, for example, for students to decline to interact with each other in a meaningful way during peer discussion, these categories nevertheless provide a useful way to characterize the way in which most students are likely to experience, and to interact with, the course material during interactive engagement activities in a lecture.

Many discussions of interactive engagement pedagogies focus primarily on the value of peer discussions. Certainly, the student-student discussions that take place during PI have been shown to be highly valuable for learning [18,19]. Yet, as discussed, although these peer conversations are a central element of Peer Instruction, it also creates opportunities for other types of interactions, for example, with the subject material through problem solving, and, central to the generation of feedback and formative assessment [20], with the lecturer through whole class discussions.

While student-student interactions have been documented in the PER literature [19,21], discussion of teacher-student interactions is much rarer, even though in the school science education literature, classroom discourse has been a central area of research for a number of decades [22–25], with both student-student discourse and teacher-student interactions gaining attention. This is surprising given the introduction and proliferation of interactive engagement techniques. It may be in part because teaching typically takes place in large lecture theatre settings in the form of traditional lectures that are didactic in nature and are therefore not expected to involve dialogue. However, discussing lectures, Bamford argues that "since language is a social activity there seems to be solid ground for looking at its organization in terms of the interaction between speakers and listeners" [9]. This is a promising approach, but Bamford and other researchers in this area focus on a traditional style of lecturing in which "one speaker speaks and a group of co-participants only listen and at most ask questions" [9]. There is some evidence that researching student-lecturer interactions in active learning lectures will be fruitful; Turpen and Finkelstein [26] studied the way that six lecturers use Peer Instruction. They found that even when lecturers followed the same basic stages for PI, there were significant differences in the type of student-lecturer interactions. For example, some lecturers allowed more time for sense making through giving opportunities to hear different opinions from students, while others sought student answers rarely. As IE pedagogies become commonplace in physics instruction, characterizing these and other types of interactions is increasingly important if we are to gain a detailed understanding of the role that interactions have in active learning lectures.

C. Theoretical approach

Our theoretical approach is influenced by activity theory, a descriptive framework for studying the contextual aspects

of different practices in a way which links the individual and social dimensions of that practice. Activity theory has its roots in the ideas of the psychologist Vygotsky. For Vygotsky, learning is mediated through signs and symbols, most notably language. This idea led Vygotsky to conclude that all learning happens initially on the social plane (i.e., through social interactions) before being transferred into the personal (internal) plane. Others, such as Leontiev and later Engeström have built on this, noting that it is not "word meaning" that should be considered the unit of analysis, but "human activity" [27]. Activity theory is therefore a framework for studying "different forms of human praxis as developmental processes" [28].

There is extensive evidence that this approach is appropriate for education research [29,30] and is therefore likely to be useful for studying lectures using interactive engagement activities that include a broad range of activity types. According to Greeno, activity theory leads to learning environments being seen "as activity systems in which learners interact with each other and with material, informational, and conceptual resources in their environment" [31].

Activity theory considers the activity system as a whole by focusing on the subject (the person undertaking the activity), the tools (the mediating device, for example, instruments, signs, and language), and the object (the intended activity) [29]. This enables the relationships between these components to be studied. In this research the subject is the students (and the lecturer), the tools are language (and also the clicker technology), and the objects are the learning activities, such as problem solving. We are particularly interested in the relationship between each of these components, in other words, the types of tools being used by the subjects in pursuit of the object goals, and we conceptualize these relationships as interactions. In this we focus particularly on the interactions between students, between student and teacher, and between the student and the material that they are studying.

III. CONTEXT AND METHODOLOGY

The University of Edinburgh introductory physics course has a history of using research-supported, innovative pedagogy and has used an inverted or "flipped" approach [32] together with Peer Instruction (PI) for a number of years [33]. Although some definitions of the flipped classroom require video lectures [34], others provide a broader approach. We follow the definition given by Abeysekera *et al.* [35], which is based on the philosophy that a flipped course is one in which the content is delivered before the lecture, thus freeing up class time to be spent on more in-depth thinking about the content. Their definition requires three components:

(1) move most information-transmission teaching out of class,
(2) use class time for learning activities that are active and social, and

(3) require students to complete pre- and/or postclass activities to fully benefit from in-class work.

In our approach the course consists of prereadings, lectures, and workshops. During the week prior to lectures on a given topic students read the course material, delivered through both electronic resources and text books, and complete a short online quiz. Approximately 85% of students complete the online quiz each week. The lectures on the topic are then predominantly focused on problem solving and discussions through the use of Peer Instruction [16]. The lecturer does not spend time delivering new content, but may provide additional explanations or demonstrations as required. While some use a strict definition of "flipped" to mean only interactive activities and no lecturing during class time, we use a wider definition of flipped based on the idea of first contact with material happening ahead of class and no introduction of completely new material during class time. We wish to make it clear, however, that some class time is used for didactic lecturing in our approach.

Turpen and Finkelstein [36] observed that different classroom practices give rise to quite distinct sets of 'classroom norms', and that students are able to perceive the differences in these norms. The norms established in Physics 1A and 1B, right from the first meeting of the class and persisting throughout the semester, are that questioning, discussion, and dialogue are a regular and integral part of the class, and that the clear expectation is that all students will participate in some way. We would hypothesize that this "norming" of questions and discussion would encourage the students to feel that they are actively part of the dialogues in class, more so than in traditional classes where occasional questions from students could be seen to be more exclusively between the instructor and the one student. It may also encourage a greater willingness in students to interject with unsolicited questions. This distinction informs our considerations of "vicarious interaction" (see later discussion).

Each lecture is approximately 50 min long. PI is implemented through the use of clickers (electronic voting systems) and follows a protocol similar to that described by Mazur [16]. For the purposes of this research, PI was considered as a five step process: (a) lecturer posing a problem, (b) students thinking individually and placing initial vote, (c) students discussing in small groups and revoting, (d) whole class discussion of the questions, and (e) lecturer summing up. A PI session always included steps (a), (b) and (e), but (c) and (d) were optional.

The course is calculus based with a gender ratio of around 80:20 males to females. Approximately half the class are majors, intending to complete a physics degree, with the remaining students being nonmajors from predominantly (but not exclusively) other STEM disciplines. The class is taught as a single section with majors and nonmajors together. It should be noted that in terms of prior

educational qualifications, the nonmajors are as well qualified as the majors: all members of the class must have satisfied the entrance requirements for the physics degree program.

Two first year (introductory level) courses were studied in this research: Physics 1A, with an enrolment of 280 students, taught in the first semester, and Physics 1B with an enrollment of 229 students, taught in the second semester. The students in Physics 1B are a subset of the students in Physics 1A. For both courses class sizes of 200 students were typical. Physics 1A covers Newtonian mechanics including kinematics and forces and Physics 1B consists of a range of different topics including introductory thermodynamics, introductory waves, applications of waves, basic physical optics, and basic crystals. The mean normalized gain on the Force Concept Inventory (FCI) for Physics 1A is 0.49 ± 0.1. Other than the content, both of these courses were taught under the same conditions. Both courses were taught in the same lecture theatre, which had a traditional layout with tiered, front facing seating and a stage area for the lecturer. Both courses were taught with the same pedagogical approach and by the same lecturer (R. K. G., who is also an author on this paper). This allows us to make direct comparisons between them. R. K. G. is an experienced physics lecturer who has been using a flipped classroom with PI for over four years. He is also actively involved in PER. However, we do not present the data here as an example of "best practice," rather our aim is to determine the scope and parameters of an active learning lecture that might be useful for future research. Such parameters could enable, for example, meaningful comparisons to be made between different teaching interventions.

A. Methodology

Although a number of tools for analyzing classroom activities have been described in the literature, it was felt that none offered the parameters necessary to answer our research questions. In particular, our sociocultural approach focusing on the interactions between student and lecturer, and between peers, is different to those described. In order to reduce any potential bias from previous work, we therefore took a grounded theory approach, deriving codes in an iterative fashion during data analysis. Our resulting framework has much in common with those described in the literature, particularly COPUS [13]. However, COPUS focuses on what the student and the lecturer are doing individually, rather than on the interactions between student and lecturer, which is our primary focus in this work as discussed in more detail later.

Video recordings of lectures from the university lecture capture system, which are made available to students through the course web-based learning environment, formed the primary data for this research. These videos show both the lecturer and a close-up of anything displayed on the AV system, such as PowerPoint slides, clicker problems, or written notes on a visualizer. Eight lectures from each course were analyzed. For Physics 1A this was the maximum number available as two were considered to be nonstandard (consisting of orientation, administration, or exam revision) and for two others the lecture capture technology failed to record the lecture. For Physics 1B there were 12 available lectures. Initially the eight lectures to be analyzed from Physics 1B were chosen at random by the first author. The lecturer on the course then confirmed that these were typical of the lecture course as a whole. Videos of lectures were analyzed using a constructivist grounded theory approach [37], utilizing a constant comparative methodology. Here initial (open) coding of data was followed by theoretical sampling in order to elaborate and refine the theoretical categories. In practice, this involved the first author assigning a code to a particular segment of the lecture and then comparing this segment to previous uses of the code. In this way detailed descriptions of how each code was used were developed. Codes were refined, added, or removed as the analysis progressed. Data collection and coding continued until categories reached saturation, i.e., until it was felt that the codes fully defined all possible activities that took place during the lectures. The final stage involved developing a theoretical understanding of how the codes were related, and the creation of categories which describe the level of interactivity. In addition to the coding of lectures, a semistructured interview was conducted by the first author (A. K. W.) with the lecturer (R. K. G.) to gain further insights into the teaching approaches used for the courses in this study.

B. Interrater reliability

A set of codes was generated as described by the first author (A. K. W.). These were then explained to a second researcher (the third author, R. D.) who coded an initial set of 6 lectures. Differences were discussed and where there was a consistent disagreement a common approach was agreed. The remaining 10 lectures were then coded, achieving an overall interrater reliability of 91%, which improved to 100% following discussion. Interrater reliability was calculated as the percentage of time for which there was agreement in coding of a segment. Disagreement was defined as a segment which was either coded differently or that had a duration that differed by more than 20 sec. However, this latter condition was rare; for the majority of codes the duration differed by no more than a few seconds and the average time difference between the two coders across all segments was 5 sec. This average difference in segment duration was used to calculate the measurement errors on the data discussed later. Cohen's kappa (which takes into account the agreement in coding that may have happened by chance) was calculated to be 0.74. However, this figure is an approximation as it assumes that all coded sections took the same length of time. In fact, many

differences in coding were only a few seconds long, and it was common for there to be agreement on the longer sections.

IV. FRAMEWORK FOR INTERACTIVE LEARNING IN LECTURES

The framework for interactive learning in lectures (FILL) was developed through our analysis of physics lectures during this research project. It should be pointed out that this framework should be considered to be at an early stage of development and more work is needed to further validate the coding system for use in other contexts.

The initial aim of the research was to develop a detailed understanding of the activities that took place during Peer Instruction, since, as mentioned earlier, this is the most commonly used interactive engagement pedagogy in undergraduate physics instruction. However, during the initial stages of the research it became apparent that interactive engagement activities were taking place even in non-PI sections of the lecture. For this reason the entire lecture was coded by the type of activity that was observed. As we are interested in the types of activities which contribute to learning, any time spent on administration at the start of the lecture was disregarded in the analysis. Two coding schemes were developed, one of which describes the type of activity that the student is involved in (see Table I) and one that refers to the stage of Peer Instruction (see Table II), for example, the students thinking, or engaging in peer discussion.

Exactly one code from each table was assigned at any given time and the codes were contiguous; i.e., when one code finished another one started. Activities that were coded include the lecturer talking and students listening (Ltalk), students discussing with each other (SS-Disc), and questions to or from the students. Two distinct types of questions were observed: the lecturer asking a question (LQ) and a student asking a question (SQ). The LQ code was used for all questions asked by the lecturer that were not rhetorical questions, including questions that were answered by a student and ones that were left unanswered (but where an answer was clearly being sought). The SQ code included both the situation where a student asks a question unprompted and when the lecturer provides an opportunity for student questions. Codes such as LQ and SQ were used for the entire interaction between lecturer and student (in other words, they covered both the question and the response to the question).

Each type of activity was additionally assigned as being interactive, vicarious interactive, or noninteractive. In general, as described, interactive activities were those which involved interactions between students, between student and lecturer, or when a student was interacting directly with the material (such as through problem solving). However, we recognize that activities such as peer discussion, which involves all the students taking an active role in the dialogue, are substantially different from activities such as questions to and from the lecturer, which directly involve only a small number of participants. For this reason we have two different categories for these different types of activity. The interactive category describes activities that are commonly thought of as interactive engagement activities. Activities which were coded as interactive included the time that students spent

TABLE I.   Codes for lecture activity. Note interactive and vicarious interactive are defined as activities in which students interact with the lecturer, with each other, and with the material. See Sec. II B for more detail.

Type of Activity	Code	Description	Interactive or Noninteractive
Lecturer talking, student listening	Ltalk	Lecturer talking to the students with no interaction.	Noninteractive
Lecturer question, student answer	LQ	Lecturer asking students a question (excluding PI questions), students answering, lecturer responding to answer. Includes lecturer questions with no student response.	Vicarious interactive
Student question, lecturer answer	SQ	Students asking a question (either unprompted or prompted by the lecturer, e.g., "does anyone have any questions?"), lecturer responding.	Vicarious interactive
Student silent thinking	S-Thinking	Students solving problems individually (normally during PI).	Interactive
Feedback on PI voting, students listening	Feedback	During a PI episode, showing or viewing the results of a clicker vote.	Interactive
Student-student discussion	SS-Disc	Students discussing a problem with each other.	Interactive

TABLE II.   Codes for Peer Instruction.

Code	Description
Clicker problem (PI)	Clicker problem being introduced
Silent (PI)	Students silently working on a PI problem
Peer discussion (PI)	Students working on a PI problem in groups
Whole class discussion (PI)	Students explaining reasoning to the whole class
Lecturer summing up (PI)	Lecturer concluding or summing up after a clicker problem
Non-PI section	All activities that are not part of Peer Instruction

actively thinking about a PI problem individually (silent) and discussing a PI problem (SS-Disc). We also introduced the category of vicarious interactive activities. This category recognizes that both lecturer questions and student questions are interactive, but that each exchange involves a small number of students. However, we also felt that even for the students not directly involved in the dialogue, an exchange of questions is significantly different in nature to listening to the lecturer talking. Questions from the lecturer result in students thinking about the answer, even if they themselves are not the one to put their hand up. Similarly if a student asks a question it is likely that other students have thought about similar questions, and will therefore be processing the exchange in a more constructive and active way. The term vicarious was chosen to highlight that the norms of the classroom invite all the other students to imagine what their contributions would be if they were active in the dialogue, in other words, they "follow along" in the discussion rather than passively waiting for it to conclude and the lecture to continue. Questions from both students and lecturer were also coded as vicarious interactive even if they did not result in a verbal response, as these are experienced as the opening "move" in a dialogic exchange.

As the coding developed a further category of "Feedback" was added to the original codes for the type of interaction. The feedback code was used specifically for a section of a PI episode in which the lecturer showed the students the voting statistics or bar chart and discussed their voting patterns. This was included in the interactive category even though there was no verbal interaction during this activity, as feedback was a direct response to the clicker vote of the students. It was, therefore, essentially part of a dialogue created through and mediated by the use of clicker technology (discussed in more detail later). Feedback did not include general discussion of the physics of the clicker problem, which was included as one of the other codes as appropriate (such as lecturer talking).

The Ltalk category (lecturer talking while students were listening) was defined as noninteractive. Time spent introducing the clicker problems was included in the Ltalk category as this was experienced in the same way as other sections of Ltalk—as the lecturer talking and the students listening with no interaction. Similarly, demonstrations used in the lectures were also coded as the lecturer talking. Although lecture demonstrations have traditionally been considered to involve more active involvement with the content than simply listening to a lecture, recent research shows that when students passively observe demonstrations they gain less understanding of the underlying concepts compared with students who are encouraged to be more actively involved, for example, by asking students to predict the demonstration outcome before seeing it [38]. As with other categories our categorization is also based on how the student experiences the activity. In this case, we observed that lecture demonstrations involve no more interaction than listening to the lecturer talking. For these reasons lecture demonstrations were not considered to be interactive.

The second coding system refers specifically to the section of Peer Instruction. A PI episode was split into five sections based on the stages defined by Mazur. Sections that were not PI were coded as Non-PI. The five steps that form PI are (a) the lecturer posing a problem, (b) students thinking individually and placing initial vote, (c) students discussing in small groups and revoting, (d) whole class discussion of the questions, and (e) lecturer summing up. Although the "lecturer summing up" section involves primarily the students listening and the lecturer talking, it is an important part of Peer Instruction as it consolidates the ideas that have been expressed by the students during the whole class discussion. It also gives students a chance to hear an explanation for the correct answer (as well as why the other answers are incorrect). During non-PI sections any of the activities in Table I can take place including the lecturer delivering a monologue or an interactive question and answer session. It may, therefore, include both interactive and noninteractive components. A lecture involving a monologue is often equated with information transfer, rather than deep conceptual understanding. However, as this is a flipped class, the students will have already encountered the material in the prelecture readings. Episodes of "lecturing" are, therefore, not about covering the content, but are more focussed, aiming to add supplementary discussion or targeted clarification of material that students are already familiar with. They are delivered "just in time" so that students are prepared and motivated to make the most of the content.

In summary, exactly two codes, one of which designates the type of activity that the students are involved in, and one of which refers to the stage of PI, were assigned at any given time. For example, if the lecturer asks a question during a whole-class discussion of a clicker problem, this

would be coded as LQ, whole class discussion. On the other hand, if students ask a question while the lecturer is explaining a concept during a lecture section of the class (i.e., not part of Peer Instruction), the entire interaction would be coded as SQ, Non-PI.

V. RESULTS

A. Interactive vs noninteractive activities

The first question that we addressed in this research is "how much time do students spend on interactivity during lectures ?" This is particularly important because teaching approaches are often described in binary terms as either "traditional" (and therefore transmissionist in approach) or as "reformed or evidence based" (and, therefore, involving interactive engagement activities). Perhaps surprisingly, there are very little published data showing how much time is spent on interactive engagement activities at the higher education level either for research informed lectures or for traditional instruction. Yet such data are vital if we are to move beyond the general assertion that active learning leads to better learning, towards developing a more detailed understanding of how active learning works, and how it can be optimized.

The bar charts in Fig. 1 shows the fraction of each lecture that was spent on interactive and vicarious interactive activities for 1A and 1B. We found that the average proportion of a lecture spent on interactive activities was $25\% \pm 2\%$ and on vicarious interactive activities was $20\% \pm 3\%$, across the two courses analyzed in this study. This is in broad agreement with research by Georgiou *et al.* who reported the time spent on interactivity (which includes both interactive and vicarious interactive, as we have defined them), during both interactive teaching courses and transmission style lecturing while studying the effect of using interactive lecture demonstrations in
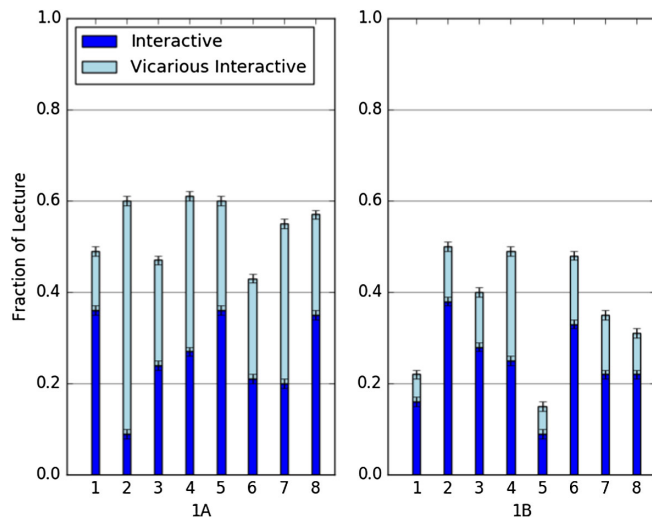
thermodynamics courses [39]. Their work found 35% and 23% interactivity for the interactive teaching streams compared to about 4% and 2% for the transmission style streams. In comparison, a paper from a cross-disciplinary study of 21 lectures in Germany observed that an average of only 9% of a lecture was spent on what they termed "advanced instruction" [40]. However advanced instruction is a broad term that included activities such as recapping material from a previous lecture, which we have defined as noninteractive. This means that the comparable figure is likely to be lower.

The TDOP protocol developed by Hora has also been used to measure the time spent by lecturers on lecturing [12], however, the results are difficult to compare to our data as TDOP codes activities in 2-min segments. Hora found that 61% of lecturers spend less than 20 min of a 50-min long lecture, lecturing (i.e., 40%). In practice, this means that he found that ten 2-min periods were coded as lecturing but as multiple codes were allowed for each segment, it is possible that other 2-min segments also contained some lecturing. The total time spent on lecturing may, therefore, be somewhat higher than 40%.

We note that in almost all of the lectures analyzed in this research, the majority of the lecture is spent on activities that are noninteractive (and often defined as passive). In fact, for some lectures in 1B, very little time is spent on any type of interactivity (though still substantially more than for a traditionally taught course). However, the average time spent on activities defined as interactive is broadly similar for each course ($26 \pm 3\%$ for 1A and $24 \pm 3\%$ for 1B). A test of statistical significance for nonparametric data (Mann Whitney test) shows that these data sets are not statistically different. A bootstrapping t test confirms this [$t(14) = 0.4$, $p = 0.695$]. However, it is interesting to note that there is a difference between the time spent on vicarious interactive activities for 1A compared to 1B, with more time being spent on lecturer and student questions in 1A on average (average of $28 \pm 4\%$ and $12 \pm 2\%$, respectively). The Mann Whitney test found that this difference is statistically significant at the 5% level ($p = 0.013$). Again, a bootstrapping t-test confirms this [$t(14) = 3.45$, $p = 0.004$]. The time spent on vicarious interactive activities in 1A varies between $13 \pm 1\%$ and $51 \pm 1\%$ and for 1B between $6 \pm 1\%$ and $24 \pm 1\%$. The possible reasons for the difference in the way that vicarious interactive activities are used between the two courses will be discussed further below.

In summary, our data show that on average a substantial fraction (55% on average) of the lecture is spent on noninteractive activities, which is significantly less than lectures which take a more traditional approach. Although there is very little work in this area in the literature, this finding is comparable to that reported by Georgiou *et al.* [39]. We also find that the time spent on interactive activities is broadly similar across the two courses, but

FIG. 1. Fractional time spent on interactive and vicarious interactive activities for 1A and 1B.

that more time is spent on vicarious interactive activities in 1A compared to 1B.

B. Classification of interactions

In order to better understand how students experience the two courses, the codes for the different types of activity were condensed into five categories that represent the five different ways that students and the lecturer interact with each other and with the material. These are as follows:

(1) Students thinking individually (normally during the first stage of PI, but also includes activities during non-PI sections, such as students being asked to draw graphs which they then display to the class).
(2) Peer discussion (students discussing a problem with each other, normally part of PI).
(3) Lecturer-student interactions (the lecturer asking students questions, students asking unprompted questions, and also the scenario where the lecturer asks "Are there any questions?" and students ask a question/make a comment).
(4) Feedback (feedback from the lecturer about the clicker votes).
(5) Lecturer talking and students listening (noninteractive).

The pie charts in Fig. 2 show the average percentage of a lecture that is spent on these five types of activity. The time spent on peer discussion is very similar for 1A and 1B (12%), as is the time spent on feedback (4.6% for both 1A and 1B). The time spent on individual thinking is also similar, with slightly more time (10% compared to 7%) being spent on this activity in 1A. However, the biggest differences are seen in the fraction of time spent on lecturer-student interactions and on lecturer talking. For both courses more time is spent on lecturer talking compared to lecturer-student interactions. For 1A 28% of the lecture is spent on lecturer-student interactions, compared to only

12% in 1B, where comparatively more time is spent on lecturer talking (64% compared to 46% for 1A).

C. Interactivity in PI vs non-PI sections

As Peer Instruction is the primary research supported IE pedagogy used in these lectures, our original hypothesis was that interactive engagement activities would take place predominantly during PI episodes, and that non-PI sections would be almost entirely noninteractive. An analysis of the proportion of time spent on both interactive learning and vicarious interactive activities for PI compared to non-PI sections of the lecture (shown in Table III) demonstrates that the situation is actually more complex than this. First, we found that the time spent on both interactive and vicarious interactive activities during PI is very similar for 1A compared to 1B (23% and 24% interactive, 9% and 8% vicarious interactive for 1A and 1B, respectively). It should be noted that we would not expect PI to be completely interactive: a PI episode consists of a range of activities, some of which we defined as noninteractive. For example, the lecturer posing the clicker problem at the start of an episode, and providing a summary of the correct answer with his explanation at the end of the episode were defined as noninteractive, as there was no obvious interaction between the lecturer and students.

For non-PI activities, we found that very little time is spent on interactive activities. In 1A this was 3% and in 1B this type of activity did not take place at all. This is not surprising as interactive activities are defined as students either working on a problem individually or discussing in small groups. Both of these are activities that are commonly part of PI, but, at least in our teaching, rarely used outside of PI. The most notable finding is that time spent on vicarious interactive activities during non-PI sections was very different for each course: 19% of the lecture is spent on vicarious activities during non-PI sections of the course for 1A and 4% on average for 1B. This implies that the non-PI sections of the lecture in 1A involved spending much more time on lecturer's students' questions. This difference almost entirely accounts for the difference seen in the average time spent on vicarious interactive activities seen between 1A and 1B, and will be explored in more detail later.
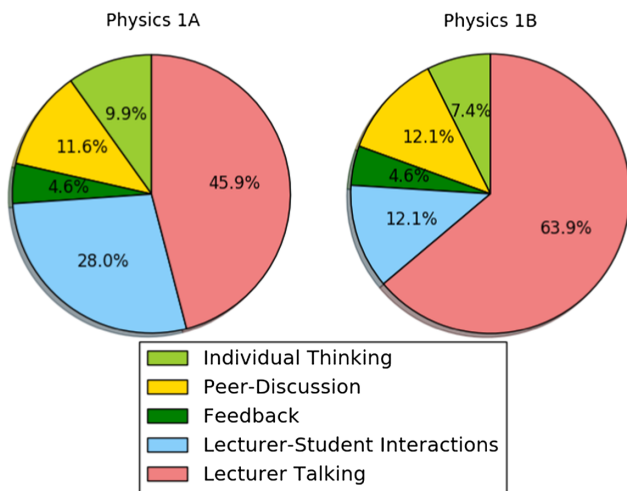

FIG. 2.   Types of interaction for 1A and 1B.

TABLE III.   Interactivity in PI compared to non-PI sections of a lecture.

% of lecture on interactive (I) and vicarious interactive (VI) activities	PI	Non-PI	Total
1A	I 23	I 3	26
	VI 9	VI 19	28
1B	I 24	I 0	24
	VI 8	VI 4	12

TABLE IV.　Average time (in minutes) spent on different stages of PI per episode (note that not all stages of PI are included in this table).

	% of lecture on PI	Total time on PI episode	Introducing Q	Silent thinking	Peer discussion	Whole class discussion	Average No. Q per lecture
1A	62 ± 0.5	$8:13 \pm 0:08$	$0:54 \pm 0:04$	$1:47 \pm 0:04$	$2:37 \pm 0:04$	$2:45 \pm 0:06$	4.5 [2-6]
1B	52 ± 0.5	$7:26 \pm 0:08$	$0:39 \pm 0:04$	$1:30 \pm 0:04$	$2:48 \pm 0:04$	$1:38 \pm 0:06$	4.3 [1-5]

1. PI active learning

As the majority of interactive learning happens during PI, it is worth examining these activities in more detail. As shown in Table II six codes were generated, five referring to elements of PI and one to non-PI sections that were coded simply as lecture. The codes for a PI episode were based on the sections of PI described by Mazur. Table IV shows the average time spent on different sections of PI during a PI episode for lectures in 1A and 1B. We found that the average number of clicker problems is similar for 1A and 1B. A similar analysis by Turpen and Finkelstein of six physics lecturers using PI in introductory classes at the University of Colorado at Boulder found that five out of the six lecturers asked more than five problems per hour of lecture [26]. The lower question count in our research may indicate that less time is spent on PI overall, or that more time is spent discussing the problem. Turpen and Finkelstein do not report how much of the lecture is spent on PI, nor the average total time spent on a PI episode. They do, however, report the length of time given for students to respond to a problem. Their implementation of PI omits the silent thinking stage, so to compare our results to theirs we need to include both the silent thinking and the peer discussion sections. Turpen and Finkelstein found that the time for students to respond to a question varied from

100 sec to just over 150 sec. In comparison, we find the average is 264 sec for 1A and 258 sec for 1B, which is substantially greater. It may be that having a silent and a peer discussion stage automatically takes more time than combining these into a single stage, or that the questions are more complex and require more thinking time.

The most significant difference between 1A and 1B is in the time spent on the whole class discussion section, with two-thirds more time spent on whole class discussion in 1A. Overall, our analysis shows that the average time spent on a PI episode is slightly higher for 1A than for 1B. This fits with the finding that in total more time is spent on interactive learning in 1A compared to 1B. However, the difference observed in the time spent on PI is not large enough to explain the total difference in time spent on interactivity found over the entire lecture course.

Figure 3 shows the fraction of each lecture spent on PI (red bars) and on non-PI activities. The most notable finding is the large variation in how PI is used for both courses. The proportion of a lecture spent on PI for 1A ranges from 27% to 92% and for 1B from 14% to 75%. The Mann Whitney test confirms that these data sets are not statistically different at the 5% level ($p = 0.96$), and a bootstrapping t test gives similar results [$t(14) = 1.57$, $p = 0.155$]. Figure 3 also shows how the non-PI time is
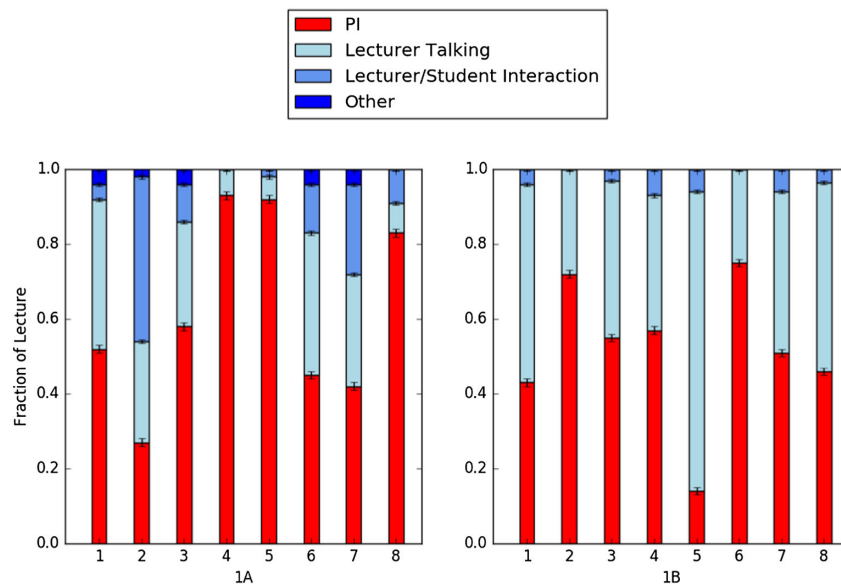

FIG. 3.　Fractional time spent on Peer Instruction. [Note that Peer Instruction consists of both interactive (38% for 1A and 33% for 1B), vicarious interactive and passive activities.]

TABLE V.   Non-PI activity summary.

% Non-PI activity	1A	1B
Lecturer talking	61 ± 0.5	92 ± 0.5
Lecturer-student interaction	34 ± 0.5	8 ± 0.5
Other	6 ± 0.5	0

used for each lecture (summarized in Table V). In most cases this time is spent predominantly on the lecturer talking (61% for 1A and 92% for 1B). Most lectures also have some form of lecturer-student interactions during this time, and more time is spent on these interactions in 1A than in 1B (34% for 1A and 8% for 1B). The "other" code includes other types of activity that happen occasionally during non-PI sections of the lecture, such as students being asked to draw a graph, or think about a problem.

As there is a large difference between the way that interactive learning (and, in particular, lecturer-student interactions) are used in non-PI sections of the course, for each course, it is worth exploring this in more detail. Figure 4 shows the fraction of the lecture spent on two types of student-lecturer interaction: lecturer questions and student questions for both PI and non-PI sections of the lecture. As the figure illustrates, the proportion of PI sections of a lecture spent on lecturer questions is similar for both courses. However in non-PI sections there are substantial differences between the two courses. Over 30% of non-PI sections of 1A lectures are spent interactively, and for the majority of that time the lecturer is asking the questions, and the students are answering them. In other words it is the lecturer who has control of the interaction. In comparison, in 1B, although much less time is spent on interactions, when they do take place, the majority are student driven and consist of students asking, and the lecturer answering questions.

In 1A it was common during non-PI sections of the lecture for the lecturer to talk through a worked example on
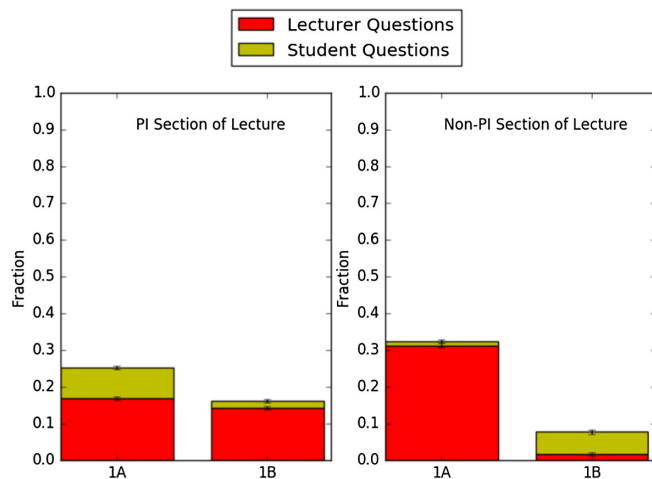

FIG. 4.   Fraction of PI and non-PI section of lecture spent on lecturer questions vs student questions for 1A and 1B.

the board and to ask for student input while he was doing so. This technique seems to be a successful way to incorporate interactivity into a portion of the lecture that might otherwise involve no interaction. In fact, the non-PI interactive learning in 1A happens predominantly during these lecturer led worked examples which are delivered in an interactive way through questions and answers. In 1B, however, this activity happens rarely, with the non-PI section of the lecture consisting of a straight didactic traditional style lecture. Where interactive engagement activities take place in non-PI sections of 1B it is through one off questions posed by the lecturer, or through student questions. A detailed analysis of the nature of these lecturer-student interactions will be presented in a subsequent paper.

VI. DISCUSSION

A. FILL framework

The framework for interactive learning in lectures presented here is at a very early stage of development, having been derived through the analysis of lectures presented in this paper. However, by taking a sociocultural approach focusing on lecturer-student interactions, it provides a new way to characterize activities in IE science lectures. The very good interrater reliability between researchers (91%), obtained through minimal training, offers hope that this framework will be easy for other researchers to adopt or adapt for their own contexts. By measuring activities on a continuous temporal basis (rather than the 2 min intervals used by some other frameworks) the FILL framework provides a high degree of precision: in fact a large proportion of coded segments had durations of less than 2 min. The framework is therefore useful for characterizing types of activity, yet is simple enough to make interpretation of results straightforward.

B. Interactive vs noninteractive

Our analysis of sixteen physics lectures found that on average 55% of the lecture time is spent on noninteractive engagement activities. This means that for many lectures a large proportion of the lecture is noninteractive, spent predominantly on the lecturer talking. This observation is particularly important because in the literature courses are often defined in binary terms as being either active or passive, reformed or traditional. Yet, as our results show, in courses that involve interactive engagement pedagogies (such as Peer Instruction), and which would therefore be described as active and reformed in nature, a substantial amount of time may also be spent on noninteractive (or traditional) learning activities. In the lectures in this research, slightly more than half of the lecture time, on average, is spent on activities which would not fit into the interactive engagement category.

Rather than thinking about active or passive as a desirable or undesirable dichotomy, researchers seeking to understand how interactive engagement activities work should therefore consider both noninteractive aspects of the lecture and the way that they work with the interactive activities. For example a clicker problem may be more (or less) effective if preceded by a mini-lecture. The learning that takes place during peer discussion may be reinforced by the lecturer explaining their answer to the clicker problem at the end of a PI session (and, therefore, modeling expert thinking). On the other hand, if a lecturer focuses only on a single correct way of thinking about the question, this may encourage students' reliance on the authority of the lecturer. Indeed, Schwartz and Bransford [41] have shown the value of 'time for telling' when used after exercises that prepare students to make the most from the lecture—in this case through analyzing contrasting cases. Peer Instruction may work in a similar way: students who are given the opportunity to think about and discuss a problem are more likely to benefit from a post-problem explanation from the lecturer. We are not aware of any research indicating the optimal amount of time that should be spent on these types of activities, although Dancy and Henderson [1] estimate, based on the timing suggested by Mazur [16], that a strategy such as Peer Instruction (PI) would result in approximately 20% of class time being spent on student activities, questions, and discussion. This would be a fruitful area for future research.

C. Inter- and intra-course variations

The results presented in this research show that the time spent on interactive activities, vicarious interactive activities, and Peer Instruction varies significantly from lecture to lecture. While it is perhaps not surprising for there to be intracourse variations, the significant differences between the two courses are particularly interesting, especially as they are taught by the same lecturer. We found that more time was spent on vicarious interactive activities (lecturer-student interactions) in 1A than in 1B (28% compared to 12%). We also found a significant difference between the time spent on vicarious interactive activities during non-PI sections of the lecture, implying that more time is spent on questions both to and from the lecturer in 1A than in 1B.

These differences may relate to the differences in the nature of each course and particularly to the depth with which the subject matter is covered. 1A is essentially an in-depth course on Newtonian mechanics including kinematics and forces. 1B, in contrast, consists of a range of different topics including introductory thermodynamics, introductory waves, applications of waves, basic physical optics, and basic crystals. Although there is an overarching theme for 1B (particles and waves), this may not become clear until the final few weeks. This course therefore consists of an introduction to many different areas of physics, rather than looking at one area in depth.

Another significant difference between the two courses relates to how new this material is for the students, and, therefore, how they experience the course. The topics covered in 1A will all, to a certain degree, be familiar to the students as they are covered at school in Scottish Higher and A-level qualifications (though in less depth). In contrast, the topics discussed in 1B are much more likely to be completely new for the students. For this reason it is perhaps both easier and more appropriate to spend a greater proportion of time on interactive elements in 1A, as discussions are more productive when students have a grasp of the basic ideas under consideration.

When asked about his approach to teaching these courses the lecturer said,

> "It's interesting, I actually consciously try not to do anything different between the courses, but I don't feel like I fully succeed. One way or another, 1B does feel (to me at least) more superficial, and I find it harder to come up with deep, conceptual PI questions. There are also more individual topics, and I do think I spend longer on transitions, setting things up, and just didactically explaining in 1B, which is consistent with the numbers you give above."

The implication here is that it is the course content which leads the lecturer to teach in a particular way.

Another significant difference between 1A and 1B is simply the order in which the courses are presented. This means that students taking 1A are new to the university and to university style studying, whereas students in 1B (who are a subset of the students in 1A) have had time to become familiar with the course. This may explain why students were more likely to ask questions during non-PI sections of the lecture in 1B compared to 1A.

Taken together, these findings highlight the importance of local conditions—factors such as students' prior experiences, the nature of the course and its content and assessment methods—in determining the most effective teaching pedagogies.

This has important implications for future research. These findings suggest that it is unlikely that it will ever be possible to precisely determine an optimum amount of time that should be spent on interactive engagement activities, because the local conditions (the background and ability of the students), and the exact nature of the course will have an impact on how a given topic should be taught. However, future research may determine approximate proportions of interactive and noninteractive teaching that are most likely to lead to good learning outcomes for particular types of courses. We also suggest that researchers should take care to specify the type of content and other relevant characteristics of the course being taught as this may affect the type of pedagogy that is appropriate.

D. Other types of interaction

In addition to verbal interactions, two other distinct "modes" of interaction were found:

(1) Nonverbal interactions. As Bamford [9] discusses, nonverbal interactions can play an important role in lecture theatre interactions. In our analysis of lectures we observed, for example, the lecturer asking for a show of hands, and laughter in response to a comment. Perhaps the most interesting example of nonverbal interactions were the occasions where students were asked to sketch a graph and to then display their drawing for the lecturer to see. Sometimes the answers were then used as possible options on a clicker problem.

(2) Technology mediated interactions. One of the difficulties of creating opportunities for interactions, particularly lecturer-student interactions, in lectures is the large number of students in the class. This can, at least partly, be overcome by the use of clicker technology. Clickers give students the opportunity to respond to a question, giving an indication to themselves, to each other, and to the lecturer of their level of understanding [20]. However, the interaction does not end there: after the vote, the lecturer shows the results to the class. In doing so he is "completing the feedback loop," continuing a dialogue with the students that has been created by posing, thinking about, and answering the clicker problem. It is for this reason that the feedback category was created in the coding system, and defined as interactive. We felt that this section of the lecture constitutes an important element of a dialogue that is mediated through the clicker technology and cannot therefore be described simply as lecturer talking. We also noticed that some students choose an invalid clicker option (such as option six in a four choice multiple choice question). It is not clear why they do this, but we speculate that this is a way for students to communicate something—perhaps that they are unsure of the answer, or that they are bored with the voting system.

These modes play an important role in the way that interactions are perceived and experienced in the lecture theatre. Although a more detailed analysis is beyond the scope of this paper, they should not be ignored in a complete description of interactions in the lecture theatre setting.

VII. CONCLUSIONS

In this paper sixteen lectures from two introductory physics courses (1A and 1B) taught by the same lecturer at the University of Edinburgh were analyzed. The aim of the analysis was to generate a detailed characterization of the interactions, as experienced by the students, that take place during flipped, interactive engagement lectures using Peer Instruction. Through this analysis a framework for interactive learning in lectures (FILL) was developed, which shows potential for characterizing lectures which use interactive engagement pedagogies. The FILL framework codes interactions on a continuous basis throughout the lecture resulting in a precise representation of activities that have taken place. Our high interrater reliability indicates that it is easy to implement and that lengthy training is not required, although we acknowledge that further work is required to assess its reliability and validity.

Our analysis found that lecture activities could be thought of as interactive, vicarious interactive (lecturer-student interactions), or noninteractive and that, on average, 25% of a lecture was spent on interactive engagement activities and 20% on vicarious interactive activities.

Interesting differences were observed between the two lecture courses studied. We found that there was a statistically significant difference in the proportion of time spent on vicarious interactive learning for each course (28% for 1A and 12% for 1B), but no statistically significant difference in the time spent on interactive activities or on Peer Instruction. We also found that the time spent on vicarious interactive learning in non-PI sections of the lecture was substantially different, with 1A involving substantially more lecturer-student interactions than 1B (34% for 1A and 8% for 1B). These differences in the way that interactive engagement activities were used between the two courses are attributed to two factors: (i) the different type of content covered in the lectures [particularly that 1A covered one topic (Newtonian mechanics) in depth whereas 1B was a general introduction to a range of topics] and (ii) how new the material was for the students (particularly that 1A built on concepts that were familiar whereas much of the material in 1B was new). This implies that the amount of time spent on interactive learning is dependent on course content, but also on local factors such as student prior knowledge, student abilities, and course structure.

Although it is not the exact percentages that are of interest *per se*, as this will vary from course to course, these findings do lead to more general conclusions. For example, we found that for many lectures at least half of the time was spent in activities that are defined as noninteractive, even though Peer Instruction was used in every lecture. We therefore conclude that describing a lecture course in binary terms as either interactive or noninteractive (or the more commonly used "active or passive") is misleading. Rather, we believe that both interactive and noninteractive components have their place in lectures and hope that future work will focus on the interplay between these components to determine how these can work optimally together.

[1] M. Dancy and C. Henderson, Framework for articulating instructional practices and conceptions, Phys. Rev. ST Phys. Educ. Res. **3,** 010103 (2007).

[2] 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).

[3] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, Proc. Natl. Acad. Sci. U.S.A. **111,** 8410 (2014).

[4] M. Dancy and C. Henderson, Pedagogical practices and instructional change of physics faculty, Am. J. Phys. **78,** 1056 (2010).

[5] M. T. Hora, Toward a Descriptive Science of Teaching: How the TDOP Illuminates the Multidimensional Nature of Active Learning in Postsecondary Classrooms, Sci. Educ. **99,** 783 (2015).

[6] C. H. Crouch and E. Mazur, Peer instruction: Ten years of experience and results, Am. J. Phys. **69,** 970 (2001).

[7] S. V. Franklin, E. C. Sayre, and J. W. Clark, Traditionally taught students learn; actively engaged students remember, Am. J. Phys. **82,** 798 (2014).

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

[9] J. Bamford, Interactivity in academic lectures: The role of questions and answers, in *Dialogue within Discourse Communities* (Niemeyer Verlag, Tubingen, 2005), pp. 123–146.

[10] R. Murphy and N. Sharma, The inverted classroom in a large enrolment introductory physics course: A case study, in Proceedings of the Higher Education Research Conference - STEM (Higher Education Academy, London, 2010).

[11] 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).

[12] M. T. Hora, A. Oleson, and J. J. Ferrare, *Teaching Dimensions Observation Protocol (TDOP) User's Manual* (Wisconsin Center for Education Research, Madison, 2013).

[13] M. K. Smith, F. H. M. Jones, S. L. Gilbert, and C. E. Wieman, The Classroom Observation Protocol for Undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices, CBE Life Sci. Educ. **12,** 618 (2013).

[14] G. M. Novak, A. Gavrini, W. Christian, and E. Patterson, *Just-In-Time Teaching: Blending Active Learning with Web Technology* (Prentice Hall, Upper Saddle River, NJ, 1999).

[15] D. R. Sokoloff and R. K. Thornton, Using interactive lecture demonstrations to create an active learning environment, in *The Changing Role of Physics Departments in Modern Universities* (AIP Publishing, College Park, 1997), Vol. 399, pp. 1061–1074.

[16] E. Mazur and R. C. Hilborn, Peer Instruction: A User's Manual, Phys. Today **50,** No. 4, 68 (1997).

[17] M. T. H. Chi, Active-constructive-interactive: A conceptual framework for differentiating learning activities, Topics Cogn. Sci. **1,** 73 (2009).

[18] L. A. Barth-Cohen, M. K. Smith, D. K. Capps, J. D. Lewin, J. T. Shemwell, and M. R. Stetzer, What are Middle School Students Talking About During Clicker Questions? Characterizing Small-Group Conversations Mediated by Classroom Response Systems, J. Sci. Educ. Technol. **25,** 50 (2016).

[19] A. K. Wood, R. K. Galloway, J. Hardy, and C. M. Sinclair, Analyzing learning during Peer Instruction dialogues: A resource activation framework, Phys. Rev. ST Phys. Educ. Res. **10,** 020107 (2014).

[20] I. D. Beatty and W. J. Gerace, Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology, J. Sci. Educ. Technol. **18,** 146 (2009).

[21] J. Tuminaro and E. F. Redish, Elements of a cognitive model of physics problem solving: Epistemic games, Phys. Rev. ST Phys. Educ. Res. **3,** 020101 (2007).

[22] D. Barnes and F. Todd, *Communication and Learning in Small Groups* (Routledge and Kegan Paul plc., London, 1977).

[23] J. L. Lemke, *Talking Science: Language, Learning, and Values* (Ablex Pub., New York, 1990), Vol. 1.

[24] N. Mercer, *Words and Minds: How We Use Language to Think Together* (Routledge, New York, 2000).

[25] J. Sinclair and M. Coulthard, *Towards an Analysis of Discourse: The Language of Teachers and Pupils* (Oxford University Press, London, 1975).

[26] C. 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).

[27] J. V. Wertsch, *Voices of the Mind: Sociocultural Approach to Mediated Action* (Harvard University Press, Cambridge, MA, 1991).

[28] D. H. Jonassen and L. Rohrer-Murphy, Activity theory as a framework for designing constructivist learning environments, Educ. Technol. Res. Dev. **47,** 61 (1999).

[29] N. H. Hashim and M. L. Jones, Activity Theory: A framework for qualitative analysis, *4th International Qualitative Research Convention, September, 2007, PJ Hilton, Malaysia* (2007), http://ro.uow.edu.au/commpapers/408.

[30] E. Scanlon and K. Issroff, Activity theory and higher education: Evaluating learning technologies, J. Comput. Assist. Learn. **21,** 430 (2005).

[31] J. G. Greeno, Learning in Activity, in *The Cambridge Handbook of the Learning Sciences*, edited by R. K. Sawyer (Cambridge University Press, Cambridge, England, 2005), pp. 79–96.

[32] S. Bates and R. Galloway, The inverted classroom in a large enrolment introductory physics course: a case study. in *Proceedings of the Higher Education Research Conference - STEM* (2012).

[33] S. P Bates, K. Howie, and A. S. J Murphy, The use of electronic voting systems in large group lectures: challenges and opportunities, New Directions **2,** 8 (2006).

[34] J. L. Bishop and M. A. Verleger, The flipped classroom: A survey of the research, in *ASEE National Conference Proceedings, Atlanta, GA* (American Society for Engineering Education, 2013), https://peer.asee.org/collections/23.

[35] L. Abeysekera and P. Dawson, Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research, Higher Educ. Res. Dev. **34,** 1 (2015).

[36] C. Turpen and N. D. Finkelstein, The construction of different classroom norms during Peer Instruction: Students perceive differences, Phys. Rev. ST Phys. Educ. Res. **6,** 020123 (2010).

[37] K. Charmaz, *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*, 2nd ed., edited by D. Silverman (Sage, London, UK, 2014).

[38] C. Crouch, A. P. Fagen, J. P. Callan, and E. Mazur, Classroom demonstrations: Learning tools or entertainment? Am. J. Phys. **72,** 835 (2004).

[39] H. Georgiou and M. D. Sharma, Does using active learning in thermodynamics lectures improve students conceptual understanding and learning experiences?, Eur. J. Phys. **36,** 015020 (2015).

[40] V. Gehlen-Baum and A. Weinberger, Teaching, learning and media use in today's lectures, Comput. Hum. Behav. **37,** 171 (2014).

[41] D. L. Schwartz and J. D. Bransford, A time for telling, Cognit. Instr. **16,** 475 (1998).