Reliable classification of classroom practices using lecture recordings

George Kinnear
G.Kinnear@ed.ac.uk
@georgekinnear
Outline

• Background
• Project overview
• Results
• Future directions
Background
Classroom practices

• Freeman et al. (2014)

• Active vs traditional

“Second-generation research could also explore which aspects of instructor behavior are most important for achieving the greatest gains with active learning” (p8413)
Using COPUS

“to verify the fidelity of the instructor to their assigned/chosen approach” (Maciejewski, 2015, p191)

Fig. 2. Summarized COPUS classroom observation data.
Smith et al. (2014)

- 51 STEM courses
- 13 departments
Stains et al. (2018)

- 2008 STEM classes
  - 709 courses
  - 548 faculty
  - 25 institutions
- Cluster analysis gave 7 clusters, grouped into:
  - Didactic
  - Interactive lecture
  - Student-centred
• Developed from literature on active learning
• Observations about distinct “activities”
• Generates scores for 21 elements, grouped into:
  – practice,
  – logic development,
  – accountability,
  – apprehension reduction.

Eddy et al. (2015)
Characterizing interactive engagement activities in a flipped introductory physics class

Anne K. Wood, Anna K. Gulliver, Robert Donnelly, and Judy Hardy
Physics Education Research Group, School of Physics and Astronomy, University of Edinburgh, Edinburgh, EH9 3JZ, United Kingdom

(Rceived 17 February 2016; published 30 June 2016)

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 (LA and IB) 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 interaction with the lecture, with each other, and with the materials 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 (15%), vicarious interactive (20%) (involving questions to and from the lecturer), and noninteractive (65%). 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 lecture-student interactions (20% for LA and 15% for IB, although not on student-student interactions (15% and 12%) or on individual lecturing (30% and 75%). These results are explored in detail and the implications for future research are discussed.

DOI: 10.1103/PhysRevEducRes.12.031014.40

I. INTRODUCTION

Interactive engagement activities developed through physics education research (PER) have been widely embraced by the physics teaching community [1]. Often used synonymously 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 255 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 6800 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 PCT gain is not guaranteed. He found that the gains for IE courses ranged from 0.25 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
Wood et al. (2016)
FIG. 2. Types of interaction for 1A and 1B.

Wood et al. (2016)
Project overview
About the project

Mathematics
George Kinnear
Pamela Docherty

Physics
Ross Galloway

Veterinary Science
Jill MacKay
Susan Rhind
Steph Smith

+ Ross Anderson, Thomas Gant
Research questions

1. To what extent do FILL and PORTAAL align (and apply across disciplines)?

2. Can classroom observation be carried out reliably using lecture recordings?

3. What patterns of classroom practices are in use at the University of Edinburgh?
Comparing FILL and PORTAAL
FILL+

- Same 1-second resolution as FILL
- New codes:
  - “Class question” rather than “clicker question”
  - Separating question and response

<table>
<thead>
<tr>
<th>Interactivity</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-interactive</td>
<td>AD</td>
<td>Admin</td>
</tr>
<tr>
<td></td>
<td>LT</td>
<td>Lecturer talk</td>
</tr>
<tr>
<td>Vicarious interactive</td>
<td>LQ</td>
<td>Lecturer question</td>
</tr>
<tr>
<td></td>
<td>SR</td>
<td>Student response</td>
</tr>
<tr>
<td></td>
<td>SQ</td>
<td>Student question</td>
</tr>
<tr>
<td></td>
<td>LR</td>
<td>Lecturer response</td>
</tr>
<tr>
<td>Interactive</td>
<td>CQ</td>
<td>Class question</td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>Student thinking</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>Student discussion</td>
</tr>
<tr>
<td></td>
<td>FB</td>
<td>Feedback</td>
</tr>
</tbody>
</table>
3. Coding Example

The details of FILL+ described in the previous section will become more apparent by actually watching a lecture and seeing the ethography being applied. The following video gives you an example of how to use FILL+ to score a 10 minute clip from a recorded lecture, with running commentary on why particular codes have been chosen.

https://osf.io/vrp7m/

The original video commentary is available to watch following this to observe the transitions between state without interruption (and with the guides of the scores given in Table 2).

https://osf.io/vrp7m/

### Table 2: Example Video scores

<table>
<thead>
<tr>
<th>Time started</th>
<th>Time finished</th>
<th>Time elapsed</th>
<th>Type of interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:01:00</td>
<td>00:02:32</td>
<td>00:01:32</td>
<td>LT</td>
</tr>
<tr>
<td>00:02:54</td>
<td>00:05:22</td>
<td>00:02:28</td>
<td>SQ</td>
</tr>
<tr>
<td>00:03:49</td>
<td>00:06:55</td>
<td>00:03:06</td>
<td>ST</td>
</tr>
<tr>
<td>00:04:06</td>
<td>00:06:57</td>
<td>00:02:51</td>
<td>FD</td>
</tr>
<tr>
<td>00:06:57</td>
<td>00:09:30</td>
<td>00:02:33</td>
<td>FB</td>
</tr>
<tr>
<td>00:07:27</td>
<td>00:10:59</td>
<td>00:03:32</td>
<td>LT</td>
</tr>
<tr>
<td>00:09:59</td>
<td>00:10:01</td>
<td>00:00:02</td>
<td>FD</td>
</tr>
<tr>
<td>00:10:01</td>
<td>00:10:05</td>
<td>00:00:04</td>
<td>SQ</td>
</tr>
<tr>
<td>00:10:05</td>
<td>00:11:01</td>
<td>00:00:56</td>
<td>LR</td>
</tr>
<tr>
<td>00:11:01</td>
<td>00:11:09</td>
<td>00:00:08</td>
<td>SQ</td>
</tr>
<tr>
<td>00:11:09</td>
<td>00:11:19</td>
<td>00:00:10</td>
<td>LR</td>
</tr>
<tr>
<td>00:11:19</td>
<td>00:11:19</td>
<td>00:00:00</td>
<td>END</td>
</tr>
</tbody>
</table>

---

**FILL+ Training Manual**

Table 1: FILL+ ethography

<table>
<thead>
<tr>
<th>Interactivity Code</th>
<th>Code</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIC</td>
<td>LT</td>
<td>Learner response</td>
<td>Any question that invites the learner to respond. For example, the learner is asked to describe an experiment, answer a question, or provide an opinion.</td>
</tr>
<tr>
<td>VIC</td>
<td>SQ</td>
<td>Student question</td>
<td>Any question that the learner is asked to respond to. The question is typically asked by the teacher or another student.</td>
</tr>
<tr>
<td>VIC</td>
<td>ST</td>
<td>Student thinking</td>
<td>Any question that the learner is asked to think about. The question is typically asked by the teacher or another student.</td>
</tr>
<tr>
<td>VIC</td>
<td>FD</td>
<td>Feedback</td>
<td>Any question that provides feedback to the learner. The question is typically asked by the teacher or another student.</td>
</tr>
</tbody>
</table>

The original video commentary is available to watch following this to observe the transitions between state without interruption (and with the guides of the scores given in Table 2).

https://osf.io/vrp7m/
# Data

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Course/lecturer combinations</th>
<th>Number of lectures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biology</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Chemistry</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Mathematics</td>
<td>21</td>
<td>108</td>
</tr>
<tr>
<td>Physics</td>
<td>9</td>
<td>60</td>
</tr>
<tr>
<td>Vet Science</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td><strong>43</strong></td>
<td><strong>234</strong></td>
</tr>
</tbody>
</table>
Reliability (I)

• Three coders
• Iterative approach:
• Carried out at start, middle, end
Reliability (II)
Results
Reliability

- Three coders by end of summer:

<table>
<thead>
<tr>
<th>Measure</th>
<th>Percent agreement</th>
<th>Krippendorff’s Alpha</th>
<th>AC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-rater</td>
<td>95.7</td>
<td>0.852</td>
<td>0.956</td>
</tr>
<tr>
<td>Intra-rater</td>
<td>96.5</td>
<td>0.849</td>
<td>0.965</td>
</tr>
</tbody>
</table>
# Reliability

- Three novice coders:

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement with model answer</td>
<td>88%</td>
<td>93%</td>
</tr>
<tr>
<td>Krippendorff’s Alpha</td>
<td></td>
<td>0.820</td>
</tr>
<tr>
<td>AC1</td>
<td></td>
<td>0.878</td>
</tr>
</tbody>
</table>

2. Can classroom observation be carried out reliably using lecture recordings? Yes
3. What patterns of classroom practices are in use at the University of Edinburgh?
Course profiles

Kinnear et al. (2020)
Interactivity

Kinnear et al. (2020)
Cluster analysis

- UG project group
- Replicating method of Stains et al. (2018)
- Found 3 clusters (proportion of LT high/med/low)
Mathematics lectures

Kinnear et al. (2020)
Peer Instruction

Question → Thinking, voting → Feedback → Discussion → Feedback

Question → Thinking, voting → Feedback
Duration of LT

Kinnear et al. (2020)
Lecturer questions

Kinnear et al. (2020)
Future directions

Comparison with COPUS  
Questioning  
Teacher intentions
Comparison with COPUS

FILL+

COPUS

?
Do you remember what Cauchy means, for a sequence to be Cauchy?

A'C' is equal to kAC and B'C' is equal to kBC. Therefore, now what?
Lecturer questions

<table>
<thead>
<tr>
<th>Paoletti et al. (2018)</th>
<th>Kinnear et al. (2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“56 questions per 80-min lecture”</td>
<td>mean of 10.7 per 50-min session</td>
</tr>
<tr>
<td>0.7 per minute</td>
<td>0.2 per minute</td>
</tr>
</tbody>
</table>

- Class size as moderator?
- Further replication of Paoletti et al. (2018)
  - question content
  - wait time
Teacher intentions

• Teaching Practices Inventory (Wieman & Gilbert, 2014)

• Comparing this with actual practice
  – Smith et al. (2014) compared with COPUS
Conclusion

- FILL+ is a reliable (and efficient) classroom observation protocol
- It gives a wealth of data to analyse practices in detail
Thank you!
References


G.Kinnear@ed.ac.uk
@georgekinnear