From Cookbook to Authentic Research: Rethinking introductory physics lab courses

@ng_Holmes
ngholmes@cornell.edu

Rochester Institute of Technology, May 3rd, 2017
Cornell Physics Education Research Lab

We’re recruiting postdocs, grad students, and undergrad students! Contact me!

Collaborators

Stanford University
Carl Wieman
Isabella Rios
Adam Stanford-Moore
Ruqayya Toorawa

University of British Columbia
Doug Bonn
Joss Ives
James Day
Dhaneesh Khumar
Sarah Gilbert
Ido Roll

Other
Jack Olsen (University of Washington)
James Thomas (University of New Mexico)
Guiding questions

What should students be learning?

What instructional approaches improve student learning?

What are students learning?

Modified from CWSEI “three-pronged approach”
What are you trying to measure?

What variables are you going to change?

What instructional approaches improve student learning?

What should students be learning?

How are you going to measure it?

What are students learning?
What are the goals of physics lab courses?

• *Think*: List some goals of intro physics labs
• *Pair*: Discuss them with your neighbor
• *Share*: Discuss with the group
Labs target...

- Understanding scientific concepts
- Interest and motivation
- Practical skills and problem solving abilities
- Scientific habits of mind
- Understanding the nature of science and measurement

Hofstein & Lunetta (1982; 2004)
there has not been much published research on the effectiveness of laboratory curricula

× Singer SR, Hilton ML, Schweingruber HA eds. (2005)
× Singer SR, Nielsen NR, Schweingruber HA eds. (2012)
Many Lab courses target...

- Understanding scientific concepts
- Interest and motivation
- Practical skills and problem solving abilities
- Scientific habits of mind
- Understanding the nature of science and measurement
What are you trying to measure?

What variables are you going to change?

Taking the lab vs not taking the lab

Course content

Final exam (lab-related and non-lab-related questions)

How are you going to measure it?

Studying the impact of labs on reinforcing course content

Holmes *et al.* (submitted)
Students who take the lab ≠ Students who do not take the lab

Must account for selection effects

Holmes et al. (in prep)
Score on lab-reinforced questions

Score on non-lab-reinforced questions

All content covered in lecture/discussion, some further reinforced in labs
Hypothesis

Lab students

Score on lab-reinforced questions

Score on non-lab-reinforced questions

No-Lab students

Score on lab-reinforced questions

Score on non-lab-reinforced questions

>
Multi-institution study

Institution 1:
- Small, private, elite research-based institution in California

Institution 2:
- Large, public research-based institution in Northwestern US

Institution 3:
- Medium, public research-based institution in southwestern US
Multi-institution study

Features:
- 3 very different populations of students
- Varied instructional approaches
- All three shared the goal to reinforce material in the rest of the course

Labs were designed to achieve that aim (e.g. making predictions, comparing results to predictions, etc.), generally quite prescribed
Labs are not providing measurable added-value to learning course content
The Colorado Learning Attitudes about Science Survey for Experimental Physics

e.g.
• When doing an experiment, I try to understand how the experimental set up works.
• When doing a physics experiment, I don’t think much about sources of systematic error.

Scores aligned with expert responses

Labs that aim to reinforce concepts decrease student attitudes towards experimental physics.

Positive shift means attitudes & belief become more expert-like.

What should students be learning???

What instructional approaches improve student learning?

What are students learning?
Labs target...

- Understanding scientific concepts
- Interest and motivation
- Practical skills and problem solving abilities
- Scientific habits of mind
- Understanding the nature of science and measurement
Cognitive tasks in experimental physics research

(Wieman, Phys. Teach. 2015)

- Establish research goals
- Define criteria for suitable evidence
- Determine feasibility of experiment
- Experimental design
- Construction & testing of apparatus
- Analyzing data
- Evaluating results & analyzing implications
- Presenting the work
Do students experience these tasks?

Numbers:
  × 8 focus-group interviews
  × 2-8 URE students per interview

Semi-structured
  × questions about URE with comparisons to coursework

Analysis
  × Look for instances where students talk about one of the cognitive tasks

Holmes & Wieman (2016) *Phys. Rev. PER*
Do students experience these tasks?

**Yes**
(They said they were doing it)

**Mixed**
(Some said yes, some said no)

**No**
(They said they were not doing it)

Holmes & Wieman (2016) *Phys. Rev. PER*
Goals and protocol laid out
Students follow steps to obtain result predicted by theory
New experiment every week

Structured Labs

vs

Design Labs

vs

Ures

Holmes & Wieman (2016) *Phys. Rev. PER*
What cognitive tasks do they do where?

Looking at # interviews where it comes up

- If something does not come up, we cannot claim it does not happen
- If students explicitly say they do not do it, we can claim it does not happen
- Mentions in repeated interviews provide strength to claim

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td>Mixed</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Holmes & Wieman (2016) *Phys. Rev. PER*
× URE doing most of the things
× Structured labs explicitly not doing the things
× Design labs, unclear – consistent “yes” but rarely discussed

![Diagram showing the number of interviews in URE, Design labs, and Structured labs across different sections: 1. Goals, 2. Criteria, 3. Feasibility, 4. Design, 5. Testing, 6. Analyzing, 7. Evaluating, and 8. Presenting. The sections are color-coded: orange for No, white for Mixed, and purple for Yes. The number of interviews in each section is indicated by the length of the bars.](image-url)
Goal Setting:

- Goals given in URE & structured labs
- Students choose research question in design-labs

Holmes & Wieman (2016) *Phys. Rev. PER*
**Design & Testing (passionate responses):**

× In URE and design labs: students had ownership, autonomy, and time to figure things out
× Lack of this in structured labs was frustrating

Holmes & Wieman (2016) *Phys. Rev. PER*
S1: When you break a machine in the one way that the professor said, ‘Do not break the machine because they don’t make spare parts for this thing anymore.’ But then you manage to fix it anyways and then the thing starts working again, that’s good... Overcoming obstacles.
S2:
I completely agree with that. Yesterday I was struggling all day long with how to fit this one graph a certain way and I was so upset and this morning I came in early and then it magically worked and I got it to work and I was so happy and it’s carried me through the whole day.
Analyzing data, evaluating results

- Despite lots of data analysis, less evaluating – not there in the UREs
- Structured labs....

Holmes & Wieman (2016) *Phys. Rev. PER*
…having the sort of, basically the amount of freedom that research does give you, having the time and the space to step back a bit and say, ‘What can we actually learn from this?’ instead of just trying to blindly get a result.
...And then sometimes when those labs, when you don’t get the results you want, you’re tempted – because you know exactly what result you want - so it’s tempting to just massage what you’ve gotten until it looks like something like a distant relative of what you want.”
UREs offer a lot that labs can’t
Clarifying career aspirations

What a professor does

What grad student life is like

Cutting edge research experiences
But labs can offer a lot more than they do now.
What should students be learning?

What instructional approaches improve student learning?

What are students learning?
Quantitative critical thinking
Quantitative critical thinking

The process through which you decide what to believe

Especially related to “believing” evidence, data, models, etc.
Quantitative critical thinking

Make a comparison

Act on comparison

Reflect on comparison
Compare period of pendulum at different amplitudes

- Measure time for single period, $T$
- Repeat 10 times, find average, standard error
Compare period of pendulum at different amplitudes

T = 1.84 ± 0.08 s

T = 1.81 ± 0.08 s

Diff ~0.2σ
Quantitative critical thinking

Make a comparison

Act on comparison

Reflect on comparison?
What might a difference of \( \sim 0.2\sigma \) mean?
What might a difference of \( \sim 0.2\sigma \) mean?

1. The periods agree
2. The periods don’t agree
3. The uncertainty is too large
4. The uncertainty is too small
\[ \text{Diff} = \frac{T_{10^\circ} - T_{20^\circ}}{\text{Uncertainty}} \]

Small difference means values are close AND/OR uncertainty is large
Quantitative critical thinking

Make a comparison

Act on comparison

Reflect on comparison
What should they do next?

- Measure time for single period, \( T \)
- Repeat 10 times, find average, standard error

\[
\begin{align*}
T &= 1.84 \pm 0.08 \text{ s} \\
T &= 1.81 \pm 0.08 \text{ s}
\end{align*}
\]

\( \text{Diff} \sim 0.2\sigma \)
What should they do next?

1. Increase the number of trials
2. Measure more swings per trial
3. Use a photogate instead of a stopwatch
4. Measure another angle
5. Write it up, list their sources of error, then go home
What should they do next?

1. Increase the number of trials
2. Measure more swings per trial
3. Use a photogate instead of a stopwatch
4. Measure another angle
5. Write it up, list their sources of error, then go home
What should they do next?

\[ T = 1.830 \pm 0.004 \text{ s} \]
\[ T = 1.851 \pm 0.004 \text{ s} \]

10° vs 20°

Diff \sim 3.7\sigma

- Measure time, \( t \), for 20 periods
- Divide by 20 to get period, repeat, average, etc.
the opposite of the expected happened:

\[ t_{\text{upon}} > 3 \implies \text{measured values are different} \]

**Conclusion:**

The period of a pendulum does depend on the angle with the vertical in the initial position.

The algebraically derived formula for \( T = 2\pi \sqrt{\frac{l}{g}} \) of a pendulum is only valid for small angles.

Considering the results of this experiment, 20° is obviously not small enough since the angle has an effect on the period and should be somehow represented in the formula.

If you can make a precise enough measurement, you can show that the theoretical derivation of the equation of motion for a pendulum is just a good approximation, and reality is slightly more complicated.
Period as a function of angle

<table>
<thead>
<tr>
<th>Angle (degrees)</th>
<th>Period (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.3</td>
</tr>
<tr>
<td>20</td>
<td>1.4</td>
</tr>
<tr>
<td>40</td>
<td>1.5</td>
</tr>
<tr>
<td>60</td>
<td>1.6</td>
</tr>
<tr>
<td>80</td>
<td>1.7</td>
</tr>
<tr>
<td>100</td>
<td>1.8</td>
</tr>
</tbody>
</table>
What should students be learning?

What instructional approaches improve student learning?

What are students learning?

Why???
Why iterative cycles work

- Autonomy and freedom to make decisions (and mistakes)
- Feedback and support to learn from decisions
- Opportunities and time to revise and improve
- Situations where physics isn’t ‘perfect’ (deal with disagreements)

Gick & Holyoak (1980, 1983); Bransford et al. (1989); Ericsson et al. (1993); Bransford & Schwartz (1999); Kapur (2008)
Can we get all students doing this?
### Assessing comparison cycles instruction

<table>
<thead>
<tr>
<th></th>
<th>Control Group</th>
<th>Experimental Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>~150</td>
<td>~140</td>
</tr>
<tr>
<td>Time</td>
<td>Weekly 3-hour labs over two semester</td>
<td></td>
</tr>
<tr>
<td>Experiments</td>
<td>Same set of mechanics and E&amp;M activities</td>
<td></td>
</tr>
<tr>
<td>Products</td>
<td>Written lab book notes</td>
<td></td>
</tr>
<tr>
<td>Instructions to iterate/improve</td>
<td>None</td>
<td>Faded out over the course</td>
</tr>
</tbody>
</table>

Holmes NG, Wieman CE, Bonn DA (2015) *PNAS*
What fraction of students in a control group do you expect to iterate without being told to?

1. Less than 25%
2. Between 25% and 50%
3. Between 50% and 75%
4. More than 75%

Iterating to improve data

- Proposed only
- Proposed & Changed
Iterating to improve data
Evaluating model issues
What are you trying to measure?

What variables are you going to change?

What should students be learning?

What instructional approaches improve student learning?

How are you going to measure it?

What are students learning?
1. Read through their lab notes

The opposite of what expected happened:

\[ \text{measured values are different} \]

**Conclusion:**

The period of a pendulum does depend on the angle with the vertical in the initial position.

The algebraically derived formula for \( T = 2\pi \sqrt{\frac{L}{g}} \)

doesn't hold only valid for small angles.

Considering the results of this experiment, 20° is obviously not small enough since the angle has an effect on the period and should be somehow represented in the formula.

If you can make a precise enough measurement, you can show that the theoretical derivation of the equation of motion for a pendulum is just a good approximation and reality is slightly more complicated.
2. Interview them

2. Criteria
3. Feasibility
4. Design
5. Testing
6. Analyzing
7. Evaluating
8. Presenting

Section

URE
Design labs
Structured labs

Number of interviews
No
Mixed
Yes
3. Design something more efficient

- Assess critical thinking in an efficient, standardized way
- Useable by instructors in different courses at any institution
Physics Lab Inventory of Critical thinking

PLIC
PLIC context

- Case studies of two student groups completing a mass on a spring experiment

Structure

- Online (Qualtrics)
- 4 sections, 2-4 questions per section
- Scoring based on alignment to experts (like C-LASS, E-CLASS)
- Paired questions (similar to Coupled Multiple Response)
For example

Students record the time for 5 bounces of a spring with 10 different masses hanging.

They plot the period (squared) as a function of mass and get:
a) How well do you think the data fit the line?  
   (Scale 1 [very bad fit] to 5 [very good fit])

b) Which items below best support your reasoning?
   • The data are in a straight line
   • There are the same number of points above and below the line
   • The points are randomly distributed above and below the line
   • The points are close to the line
   • There are too few points with error bars crossing the line...
Generating a closed-response survey

Free response (v1, v2, ...)

- Interviews + written responses
- Use common student responses to generate closed-response options
- Refine questions

Closed response (v1, v2, ...)

- Interviews + Closed-answer responses
- Statistical analysis (distinguishability, clusters, factors, etc.)
- Refine questions & options
- Remove rarely selection options

Refine questions & options
Want to use the PLIC?

Contact me
(ngholmes@cornell.edu)

Also looking for responses from experts!
Next up...

**But how?**
- Digging into the mechanism of developing critical thinking

**So what?**
- Does critical thinking in intro physics transfer?

**For who?**
- Does critical thinking instruction differentially impact different students?
What should students be learning?

What instructional approaches improve student learning?

What are students learning?

The PLIC

- Score on lab-related questions
- Score on non-lab-related questions

Act on comparison

Reflect on comparison

Make a comparison

The PLIC

1. Goals
2. Criteria
3. Feasibility
4. Design
5. Testing
6. Analyzing
7. Evaluating
8. Presenting

Number of interviews

No Mixed Yes

<table>
<thead>
<tr>
<th>Score on lab-related questions</th>
<th>Score on non-lab-related questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What should students be learning?

What instructional approaches improve student learning?

What are students learning?

The PLIC

1. Goals
2. Criteria
3. Feasibility
4. Design
5. Testing
6. Analyzing
7. Evaluating
8. Presenting

Number of interviews

No Mixed Yes

<table>
<thead>
<tr>
<th>Score on lab-related questions</th>
<th>Score on non-lab-related questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We’re hiring postdocs and recruiting students! Contact me!

Collaborators

Stanford University
- Carl Wieman
- Isabella Rios
- Adam Stanford-Moore
- Ruqayya Toorawa

University of British Columbia
- Doug Bonn
- Joss Ives
- James Day
- Dhaneesh Khumar
- Sarah Gilbert
- Ido Roll

Other
- Jack Olsen (University of Washington)
- James Thomas (University of New Mexico)

Cornell Physics Education Research Lab

Natasha Holmes  Katherine Quinn