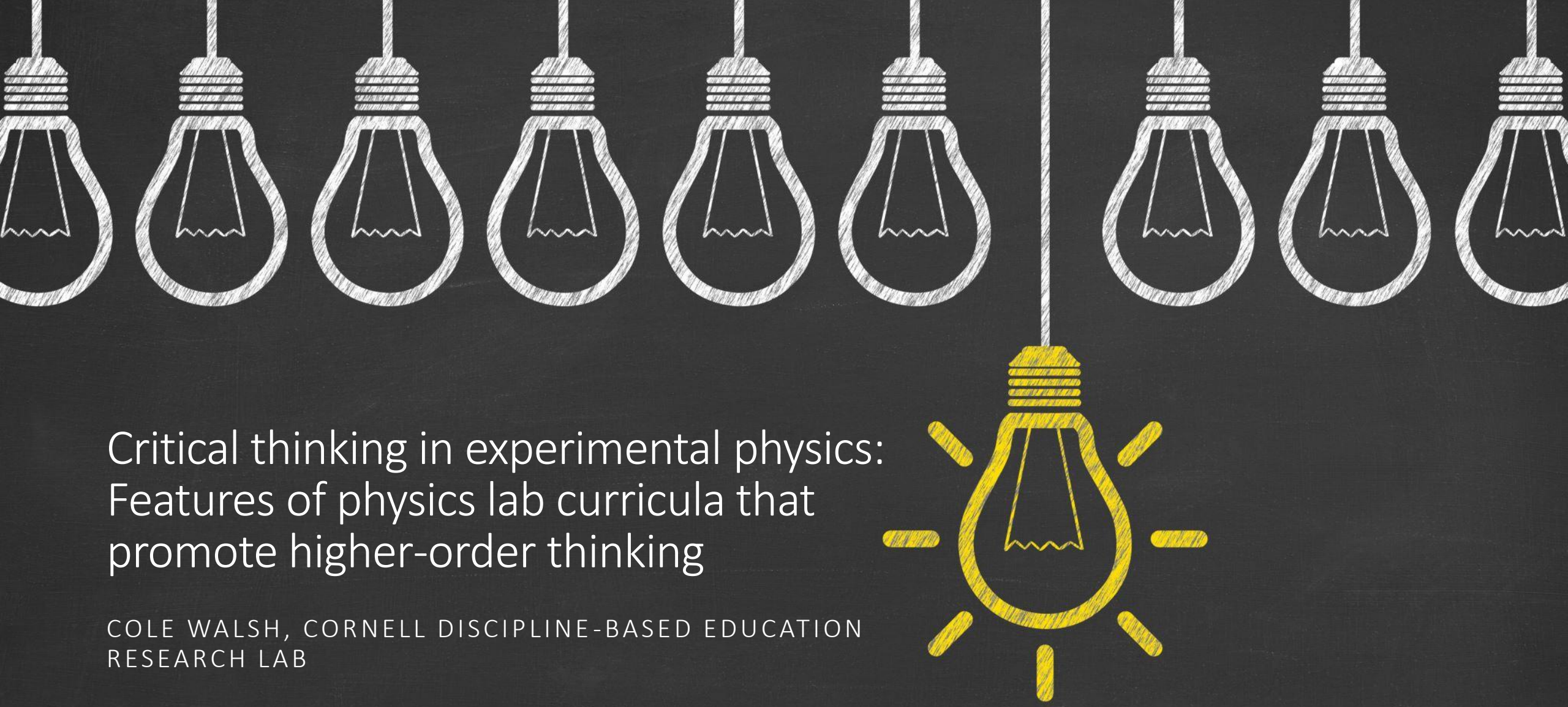


Critical thinking in experimental physics: Features of physics lab curricula that promote higher-order thinking

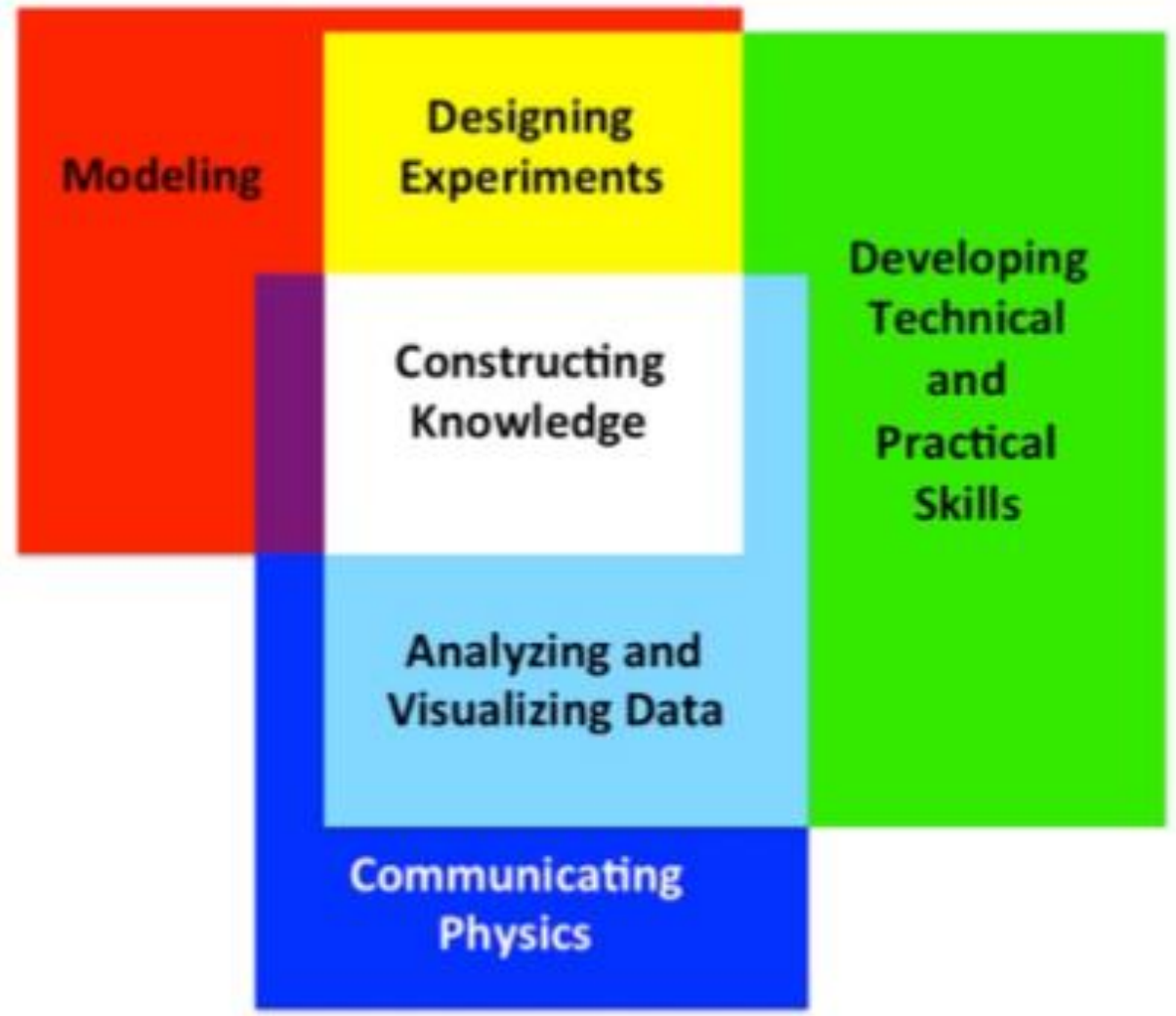
COLE WALSH, CORNELL DISCIPLINE-BASED EDUCATION RESEARCH LAB

APS APRIL MEETING 2020

What are we teaching in physics labs?

- Recent work has found:
  - no impact of lab instruction on student exam grades or conceptual understanding
  - highly structured labs negatively impact students' perception of experimental physics
  - labs designed to reinforce concepts taught in lecture negatively impact students' perception of experimental physics

HOLMES ET AL., 2017; SMITH ET AL., 2020; WILCOX AND LEWANDOWSKI, 2016, 2017



AAPT guidelines for lab instruction

Critical Thinking (CT)

Are we teaching CT in physics labs?

Can we measure students' CT skills?

The Physics Lab Inventory of Critical Thinking (PLIC)

- Critical thinking = ways in which one uses data and evidence to make decisions about what to trust and what to do.
- Developed from:
  - Student interviews,
  - Feedback from expert physicists,
  - And over 10 000 students from more that 40 schools who took open- and closed-response versions of the assessment

Summary

Group 1:
Method: 10 trials of 5 oscillations for 2 different masses and compared *k* values:
With 30 g mass: k = 4.0857 +/- 0.08067 N/m
With 50 g mass: k = 4.1872 +/- 0.06745 N/m

Group 2:
Method: 2 trials of 5 oscillations for 10 different masses and compared the data to a model equation:
With intercept set to 0: k = 4.0535 +/- 0.07005 N/m
With intercept not fixed at 0: k = 4.6274 +/-0.1587 N/m

Click here to see the intro again, here to see Group 1's lab book, here to see Group 2's initial lab book notes, and here to see Group 2's second set of lab book notes.

Which group do you think did a better job of testing the model?

- ○ Group 1
- ○ Group 2
- ○ Both the same

What features were most important for comparing the two groups?
Please select no more than 3 options.

- ☐ The number of bounces of the spring per trial
- ☐ The number of repeated trials for each mass
- ☐ The number of masses tested
- ☐ How they tested other possible variables
- ☐ How they controlled variables
- ☐ The equipment they used (i.e. stopwatches)
- ☐ How they accounted for human error
- ☐ How close the results are to what the model predicts
- ☐ The size of the uncertainty (or variability between data)
- ☐ Their analysis and calculations
- ☐ How clear, organized, or detailed their lab notes are
- ☐ Other (Please describe)

PLIC Example Page

Summary

Group 1:
Method: 10 trials of 5 oscillations for 2 different masses and compared *k* values:
With 30 g mass: k = 4.0857 +/- 0.08067 N/m
With 50 g mass: k = 4.1872 +/- 0.06745 N/m

Group 2:
Method: 2 trials of 5 oscillations for 10 different masses and compared the data to a model equation:
With intercept set to 0: k = 4.0535 +/- 0.07005 N/m
With intercept not fixed at 0: k = 4.6274 +/-0.1587 N/m

Click here to see the intro again, here to see Group 1's lab book, here to see Group 2's initial lab book notes, and here to see Group 2's second set of lab book notes.

Two hypothetical groups

Which group do you think did a better job of testing the model?

- ○ Group 1
- ○ Group 2
- ○ Both the same

What features were most important for comparing the two groups?
Please select no more than 3 options.

- ☐ The number of bounces of the spring per trial
- ☐ The number of repeated trials for each mass
- ☐ The number of masses tested
- ☐ How they tested other possible variables
- ☐ How they controlled variables
- ☐ The equipment they used (i.e. stopwatches)
- ☐ How they accounted for human error
- ☐ How close the results are to what the model predicts
- ☐ The size of the uncertainty (or variability between data)
- ☐ Their analysis and calculations
- ☐ How clear, organized, or detailed their lab notes are
- ☐ Other (Please describe)

PLIC Example Page

Summary

Group 1:
Method: 10 trials of 5 oscillations for 2 different masses and compared *k* values:
With 30 g mass: k = 4.0857 +/- 0.08067 N/m
With 50 g mass: k = 4.1872 +/- 0.06745 N/m

Group 2:
Method: 2 trials of 5 oscillations for 10 different masses and compared the data to a model equation:
With intercept set to 0: k = 4.0535 +/- 0.07005 N/m
With intercept not fixed at 0: k = 4.6274 +/-0.1587 N/m

Click here to see the intro again, here to see Group 1's lab book, here to see Group 2's initial lab book notes, and here to see Group 2's second set of lab book notes.

Use multiple choice and Likert-items to focus students' thinking

Which group do you think did a better job of testing the model?

- ○ Group 1
- ○ Group 2
- ○ Both the same

What features were most important for comparing the two groups?
Please select no more than 3 options.

- ☐ The number of bounces of the spring per trial
- ☐ The number of repeated trials for each mass
- ☐ The number of masses tested
- ☐ How they tested other possible variables
- ☐ How they controlled variables
- ☐ The equipment they used (i.e. stopwatches)
- ☐ How they accounted for human error
- ☐ How close the results are to what the model predicts
- ☐ The size of the uncertainty (or variability between data)
- ☐ Their analysis and calculations
- ☐ How clear, organized, or detailed their lab notes are
- ☐ Other (Please describe)

PLIC Example Page

Summary

Group 1:
Method: 10 trials of 5 oscillations for 2 different masses and compared *k* values:
With 30 g mass: k = 4.0857 +/- 0.08067 N/m
With 50 g mass: k = 4.1872 +/- 0.06745 N/m

Group 2:
Method: 2 trials of 5 oscillations for 10 different masses and compared the data to a model equation:
With intercept set to 0: k = 4.0535 +/- 0.07005 N/m
With intercept not fixed at 0: k = 4.6274 +/-0.1587 N/m

Click here to see the intro again, here to see Group 1's lab book, here to see Group 2's initial lab book notes, and here to see Group 2's second set of lab book notes.

Which group do you think did a better job of testing the model?

- ○ Group 1
- ○ Group 2
- ○ Both the same

What features were most important for comparing the two groups?
Please select no more than 3 options.

- ☐ The number of bounces of the spring per trial
- ☐ The number of repeated trials for each mass
- ☐ The number of masses tested
- ☐ How they tested other possible variables
- ☐ How they controlled variables
- ☐ The equipment they used (i.e. stopwatches)
- ☐ How they accounted for human error
- ☐ How close the results are to what the model predicts
- ☐ The size of the uncertainty (or variability between data)
- ☐ Their analysis and calculations
- ☐ How clear, organized, or detailed their lab notes are
- ☐ Other (Please describe)

PLIC Example Page

Interested in students' answers to multiple response questions

Overview of assessment

Page 1, Group 1

1. How well does data fit model?
2. How good was their method?
3. What should they do next?

Page 2, Group 2

4. How well does data fit model?
5. How good was their method?
6. What should they do next?

Page 3, Group 2

7. How well does data fit model?
8. Which fit should Group 2 use?
9. What should they do next?

Page 4, compare groups

10. What features were most important for comparing the two groups?

Summary

Group 1:
Method: 10 trials of 5 oscillations for 2 different masses and compared *k* values:
With 30 g mass: k = 4.0857 +/- 0.08067 N/m
With 50 g mass: k = 4.1872 +/- 0.06745 N/m

Group 2:
Method: 2 trials of 5 oscillations for 10 different masses and compared the data to a model equation:
With intercept set to 0: k = 4.0535 +/- 0.07005 N/m
With intercept not fixed at 0: k = 4.6274 +/-0.1587 N/m

Click here to see the intro again, here to see Group 1's lab book, here to see Group 2's initial lab book notes, and here to see Group 2's second set of lab book notes.

Which group do you think did a better job of testing the model?

- ○ Group 1
- ○ Group 2
- ○ Both the same

What features were most important for comparing the two groups?
Please select no more than 3 options.

- ☐ The number of bounces of the spring per trial
- ☐ The number of repeated trials for each mass
- ☐ The number of masses tested
- ☐ How they tested other possible variables
- ☐ How they controlled variables
- ☐ The equipment they used (i.e. stopwatches)
- ☐ How they accounted for human error
- ☐ How close the results are to what the model predicts
- ☐ The size of the uncertainty (or variability between data)
- ☐ Their analysis and calculations
- ☐ How clear, organized, or detailed their lab notes are
- ☐ Other (Please describe)

Page 4

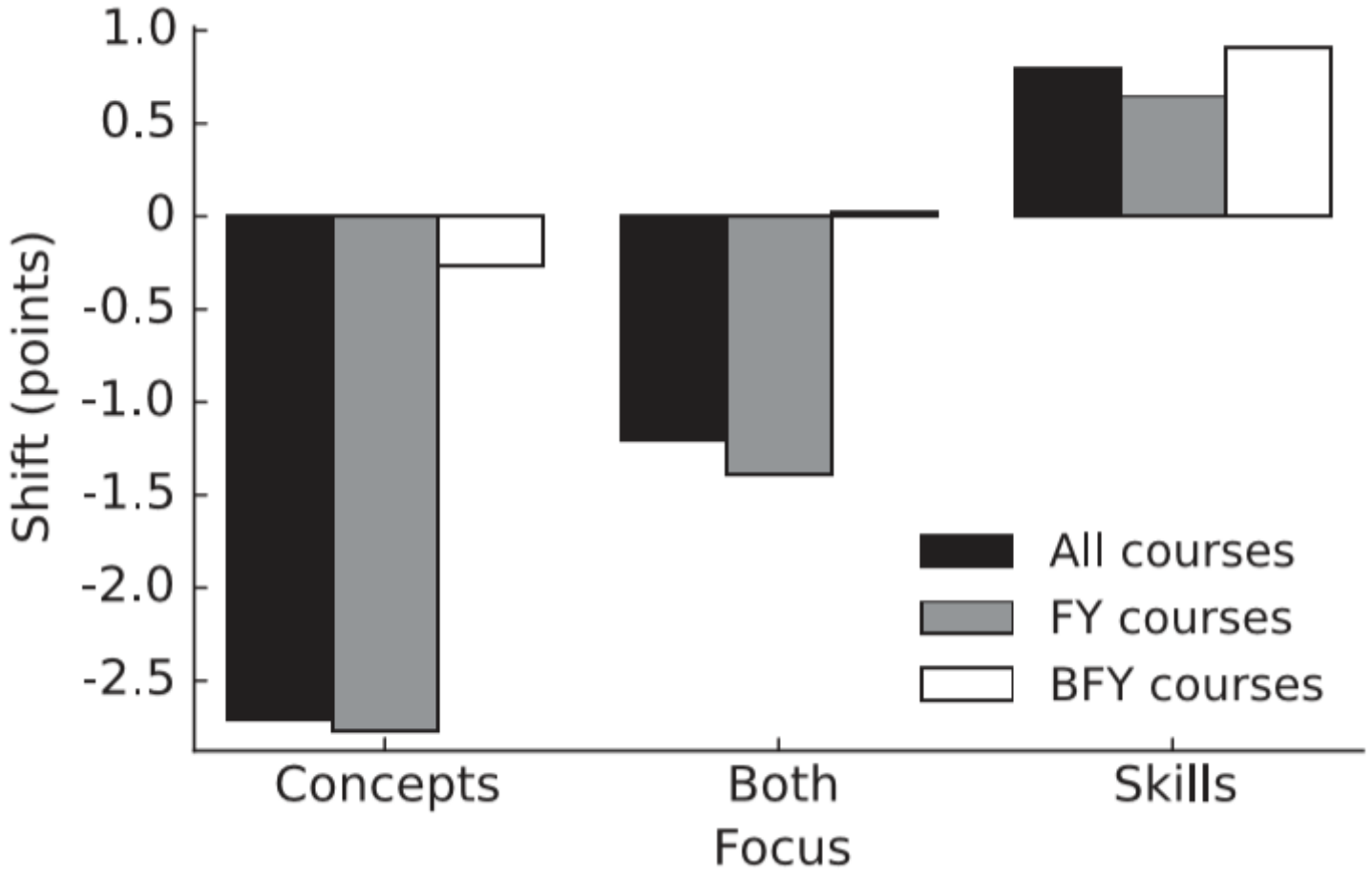
Q10

Critical Thinking

- Critical thinking = ways in which one uses data and evidence to make decisions about what to trust and what to do.
- In the context of experimental physics:
  - comparing, evaluating, and interpreting models and data,
  - evaluating methods,
  - and deciding how to proceed in an investigation.
- Students receive a score between 0 and 1 for each question based on agreement with average expert
  - Total scores range from 0 to 10

ENNIS (1989); WILLINGHAM (2007)

Are we teaching CT in physics labs?

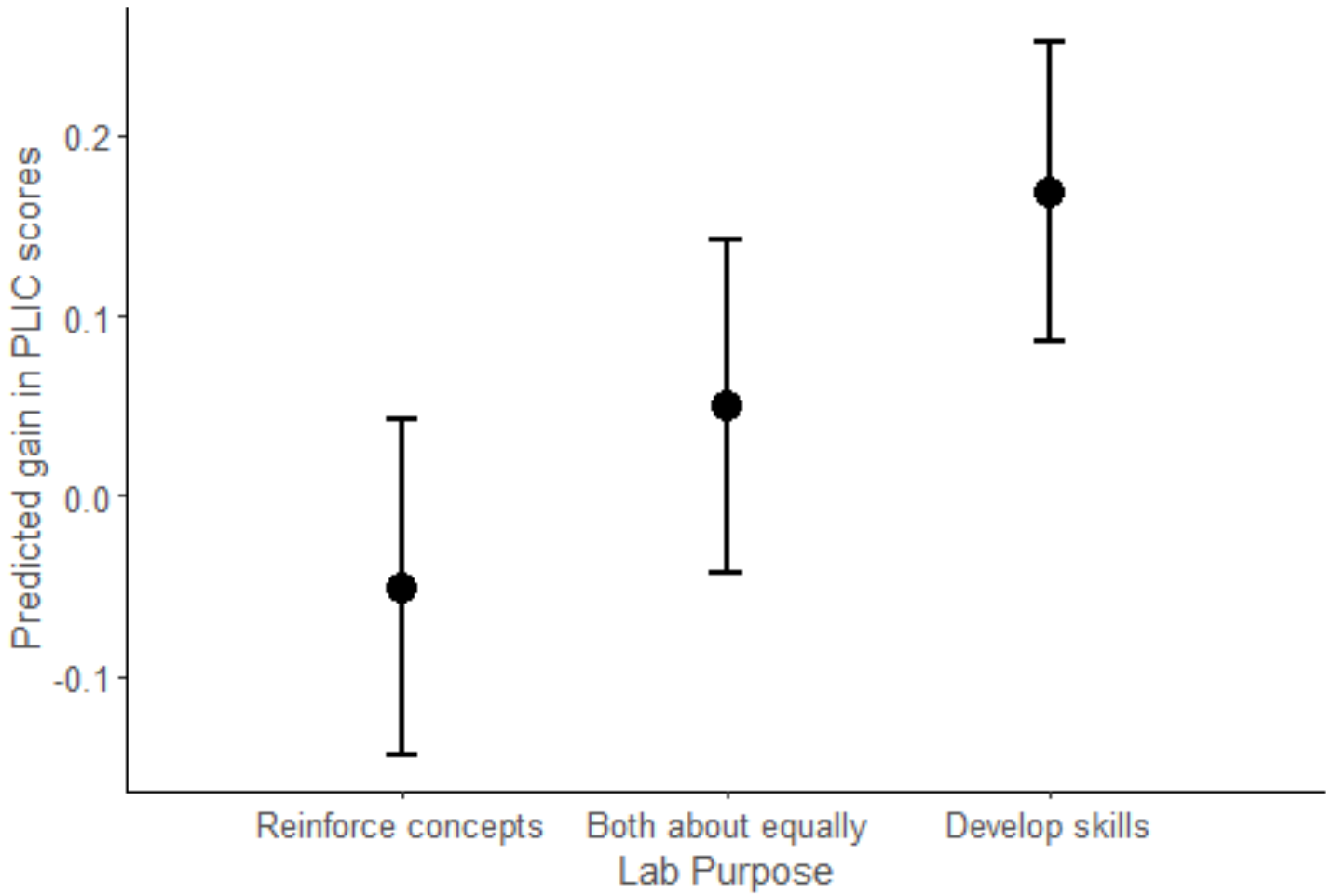


Previous work on students' perceptions of experimental physics

WILCOX AND LEWANDOWSKI, 2017

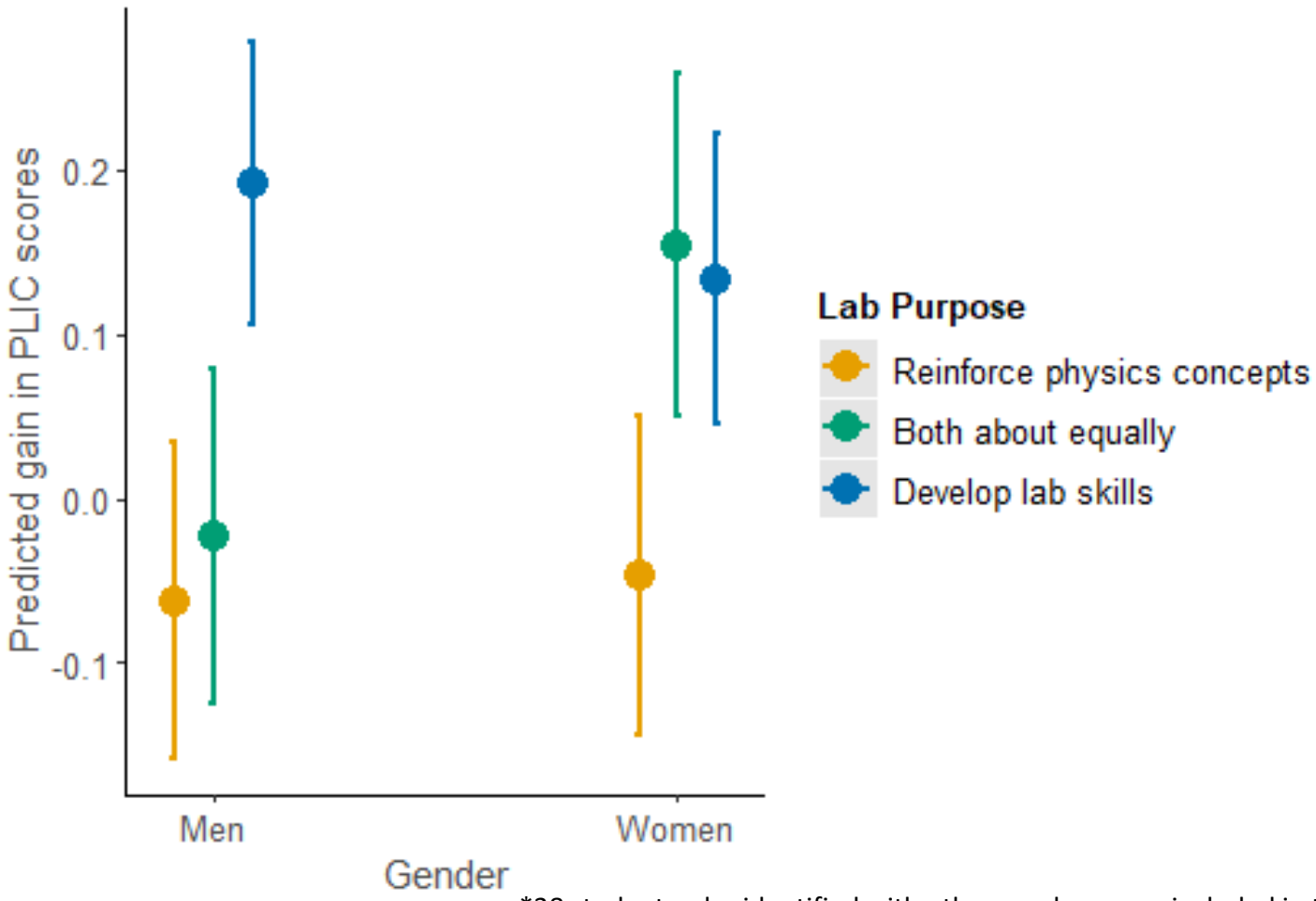
Data used in this study

- 4,174 students
- 22 schools
- 50 unique classes
  - 16 classes had labs designed to reinforce concepts from lectures
  - 21 had labs designed to teach lab skills
  - 13 had labs designed to do a combination of both



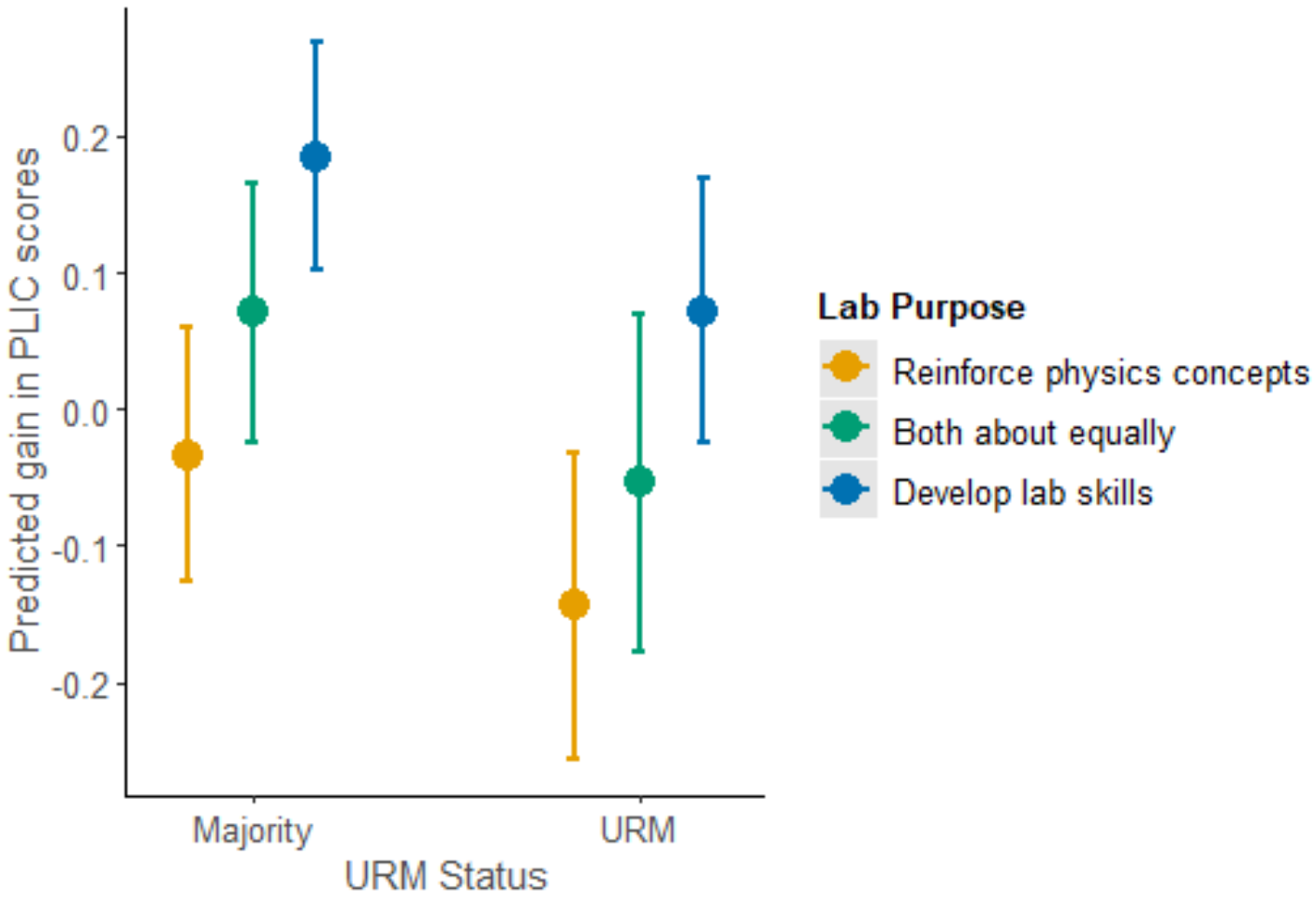
Students trained in labs designed to teach lab skills improve more on the PLIC.

Does everybody benefit from skills-based labs?



Both men and women see roughly the same benefit from skills-based labs

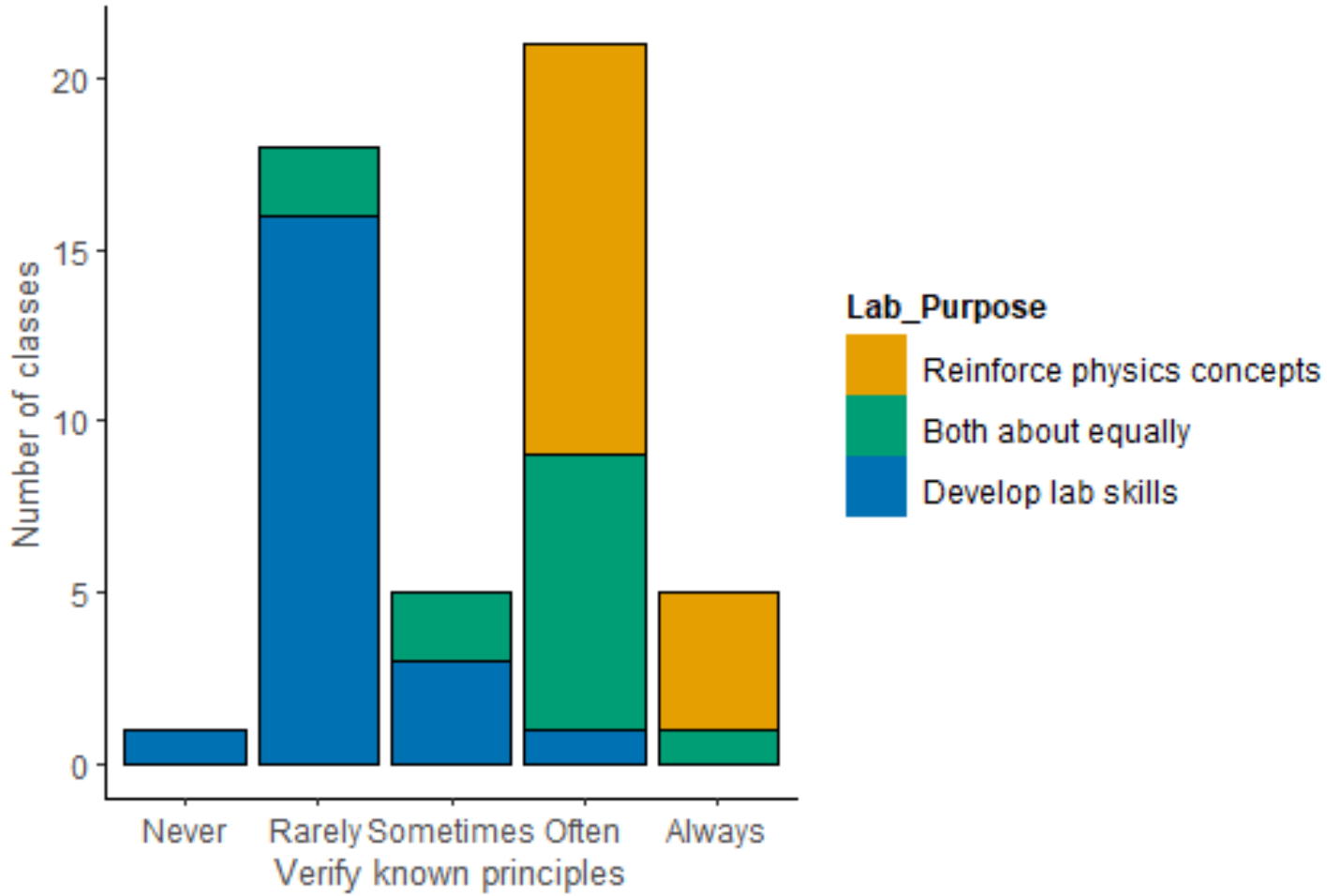
*38 students who identified with other genders were included in the analyses, but results are not presented in this graph as the error bars are too large to make any meaningful interpretations



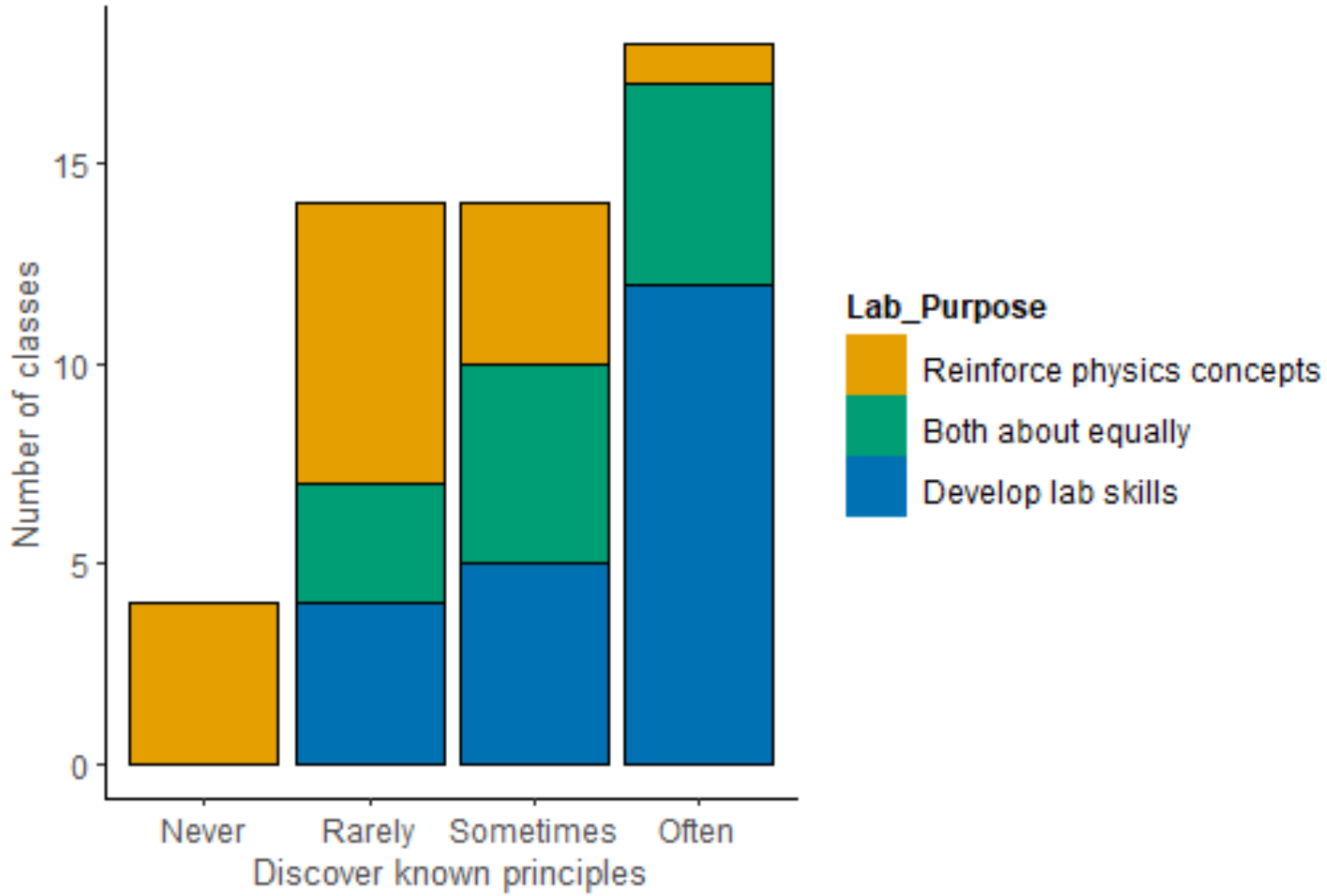
Students from traditionally represented (Majority) and underrepresented races/ethnicities (URM) also see comparable benefit from skills-based labs

Is it the purpose of the lab though?

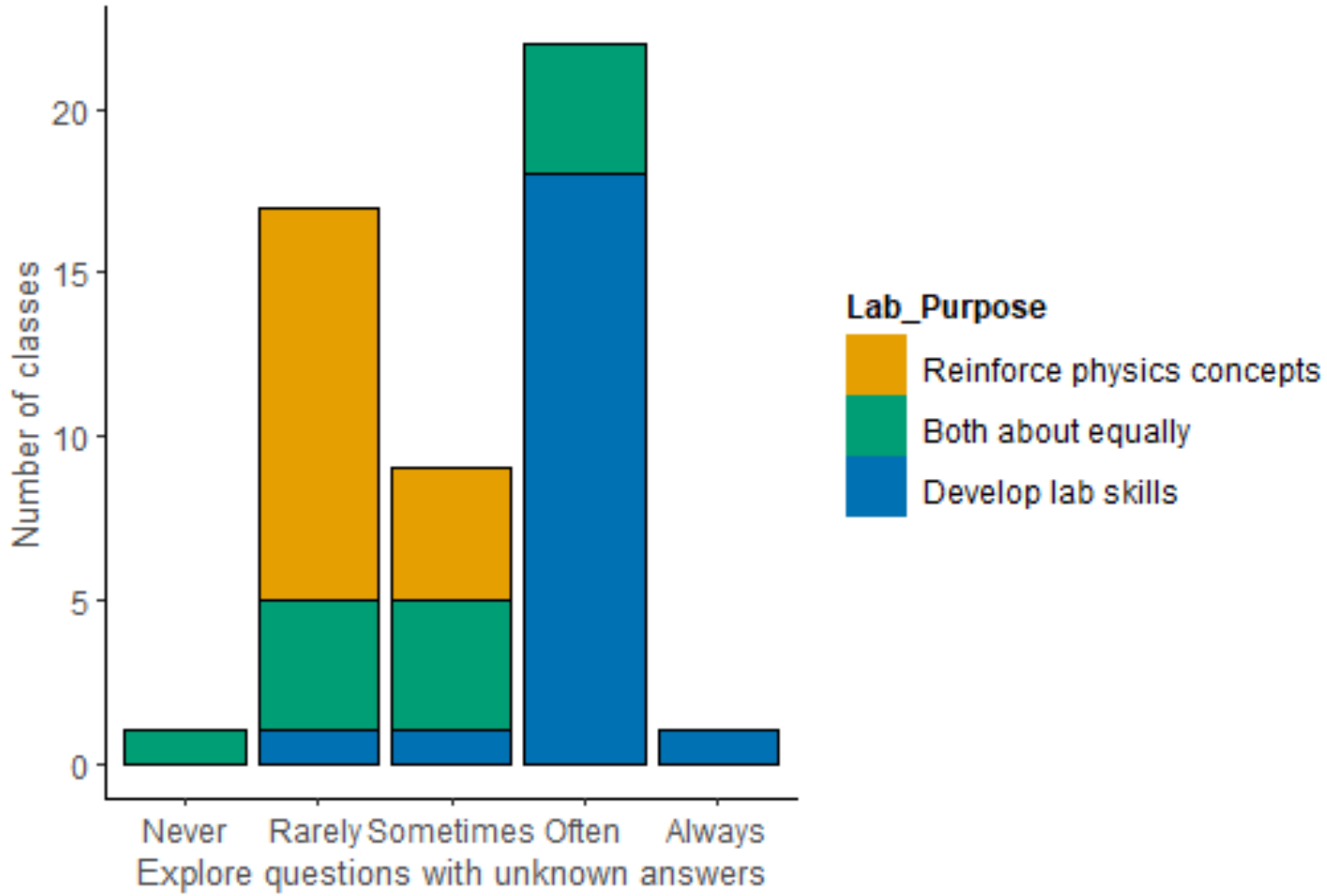
- We asked instructors many other questions before administering the PLIC including:
  - How often do students…
    - verify known principles through experimental tests?
    - discover known physical principles through experimentation?
    - explore questions to which the answer is unknown to the student?



Is it the purpose of the lab though?



Is it the purpose of the lab though?



Is it the purpose of the lab though?

Conclusions and future work

- Labs designed to develop lab skills (or have students explore questions with unknown answers) improved students' performance on the PLIC the most
- Effect was constant across students of different genders/races/ethnicities
- Gains were small
  - Consistent with other tests of general CT ability
  - Only looked at one aspect of labs
- Future:
  - Examine effect of giving students agency on performance
  - Look at correlation with particular lab activities

Thank You!

Acknowledgements

- Heather Lewandowski
- John Aiken
- CDER Lab
- National Science Foundation
  (Grant No. 1611482)