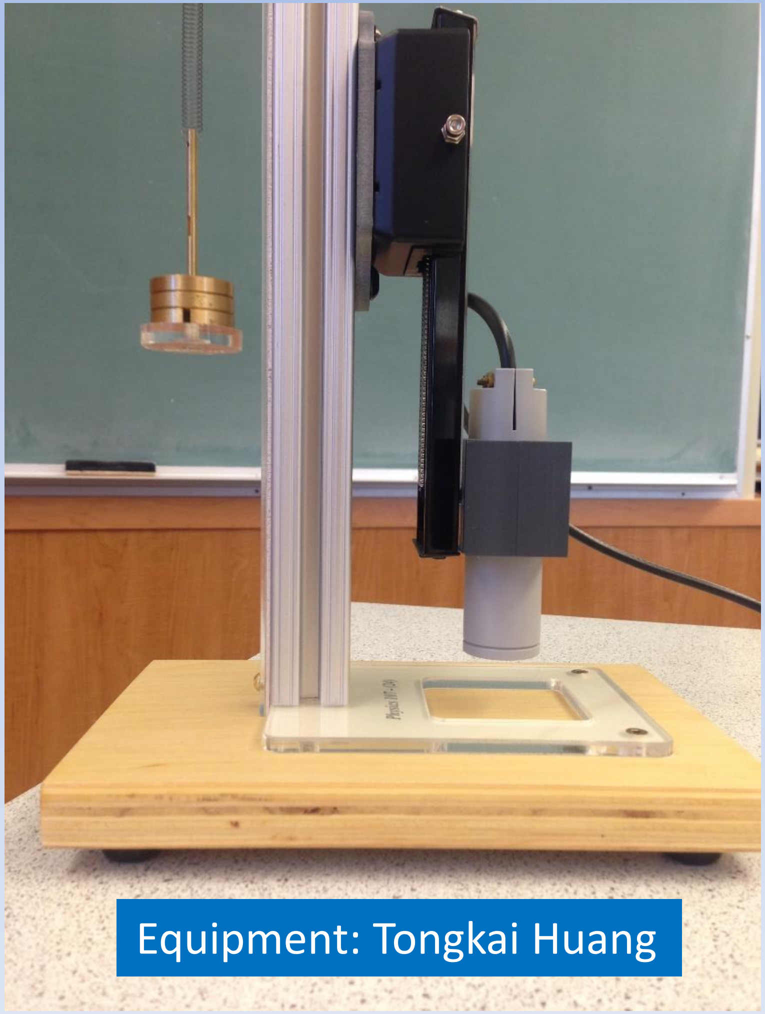
Making Comparisons: A Strategy for Teaching Scientific Reasoning

Natasha Holmes
Carl Wieman
James Day
Dhaneesh Kumar
Ido Roll
Joss Ives
Linda Strubbe

…and all of the TAs and students in UBC first year physics



Outline

- The Structured Quantitative Inquiry SQILabs
- Data showing improved quantitative reasoning
- Factors affecting the improved behavior: cuing

What is a first year physics lab for?

Support the learning of concepts covered in lectures?

What is a first year physics lab for?

Support the learning of concepts covered in lectures?

But, there are many, often hidden, goals and tasks…

Learn to use new apparatus
Learn data handling methods
Keep a lab notebook
Making formal write-ups
Oral Presentations
Measurement uncertainty
Propagation of uncertainty
Learn to use data acquisition software
Try to debug non-functional apparatus
Figuring out how to get grades
Learning time management

Learn to use data analysis software
Learn a programming language
Learn English
Develop scientific reasoning
Learn the 'Scientific Method'
Learn experimental design
Proper formatting of graphs and tables

Cognitive overload!

Some approaches to lab cognitive load

- ***Traditional lab - provide detailed 'cook book' instructions for experiments***

  - inauthentic, ineffective at teaching physics concepts

  Wieman and Holmes, AAPT 83, 972 (2015)

- ***Inquiry-based learning, studio physics***

  - careful integration of lectures/labs/tutorials – eg. ISLE labs

  Etkina and Van Heuvelen, in Research Based Reform of University Physics, AAPT, 2016

- ***Structured Quantitative Inquiry (SQILabs)***

  - drop physics goals, to concentrate on data and reasoning

  Holmes, Wieman, & Bonn, PNAS 112, 11199-*11204, 2016*

The ‘Scientific Method’ ?

Formulate hypothesis

↓

Perform experiment

↓

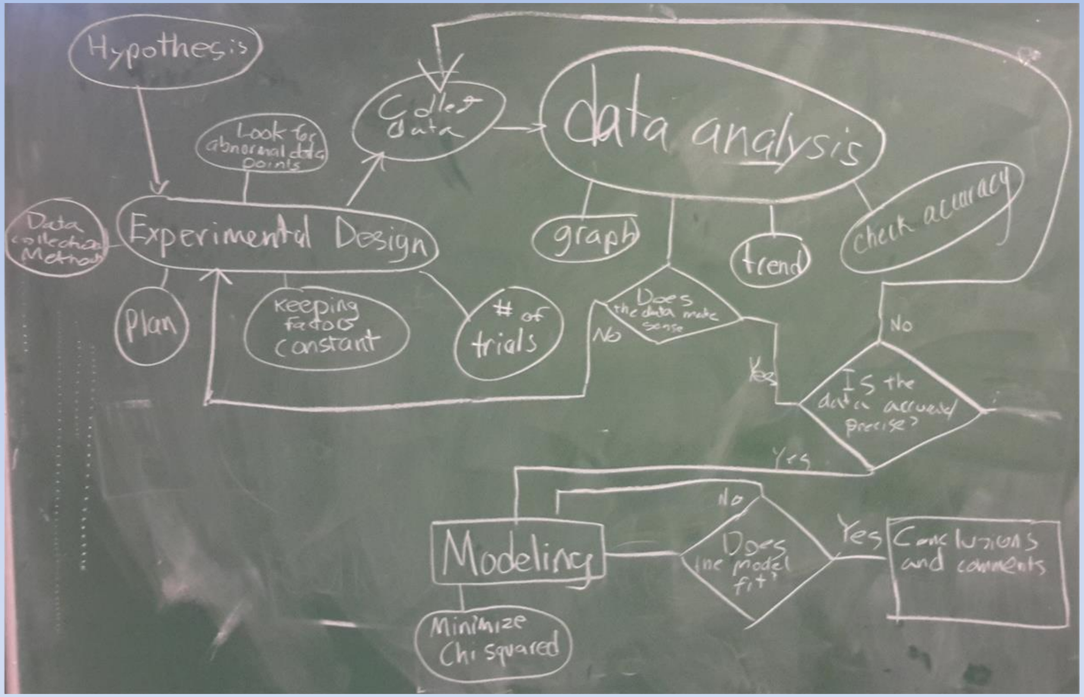
Analyze data

↓

Draw conclusion (too often, this means low level comments on bad data, poorly fit to a model, with insufficient time to succeed)

The ‘Scientific Method’ ?

Formulate hypothesis



Perform experiment

Analyze data

Draw conclusion (too often, this means low level comments on bad data, with insufficient time to succeed)

Goals and Structure of SQILabs

Goals

- *Develop a functional understanding of measurement uncertainty*
- *Learn a set of broadly applicable data-handling skills*
- *Develop expert habits of mind and scientific reasoning*

Structure

Quantitative comparisons and iteration/improvement

- Plan measurements
- Do measurements
- Make a comparison
- Reflect on comparison
- Plan an improvement

Quantitative comparison tools *t′* and χ^2

$$t' = \frac{A - B}{\sqrt{\delta_A^2 + \delta_B^2}}$$

Comparing measurements A and B

$$\chi_w^2 = \frac{1}{N - p} \sum_{i=1}^{N} \left(\frac{y_i - f(x_i)}{\delta y_i} \right)^2$$

Comparing data (y_i) and model $f(x_i)$

Plus graphical tools: histograms, scatter plots, semi-log, log-log, residuals

Bringing Comparisons, Uncertainty and Iterative Improvement together

$t'<1$	**Possible agreement?** Iterate to improve measurement; reduce uncertainty, hidden disagreement?	$\chi^2<1$
$1<t'<3$	**Tension?** Iterate to improve measurement; reduce uncertainty	$1<\chi^2<9$
$3<t'$	**Possible disagreement?** Iterate to improve measurement; remove systematic error, evaluate model	$9<\chi^2$

$$t' = \frac{A-B}{\sqrt{\delta_A^2+\delta_B^2}}$$

$$\chi_w^2 = \frac{1}{N-p}\sum_{i=1}^{N}\left(\frac{y_i - f(x_i)}{\delta y_i}\right)^2$$

Pendulum for Pros

Part II - 20 + 20 minutes (plan/measure + analyze/discuss)
The goal is to see if the period of a pendulum depends on the amplitude of the swing.
First, write down a plan for a high-precision measurement of the period at 10 degrees and at 20 degrees. Allow for roughly 15 minutes to do the measurements.
Compare your results at 10 and 20 degrees.
Part III - 20 + 20 minutes (plan/measure + analyze/discuss)
Based on your result above, write a plan for improving the quality of your measurements.
Discuss this plan with other groups at your table.
Do revised measurements and analysis.

Part IV - Keep repeating this cycle of comparing and improving, until you are confident that you understand whether or not there is amplitude-dependence in the period.

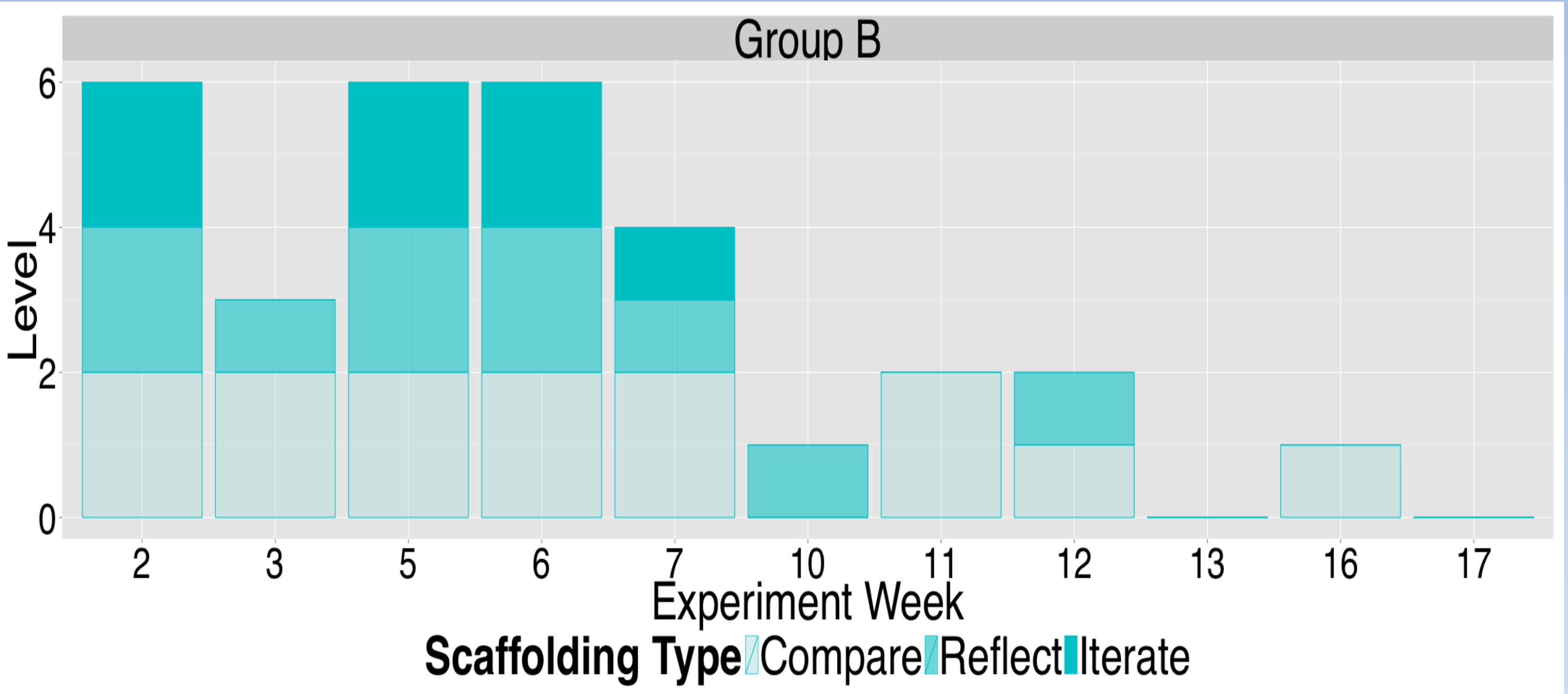
Marking Scheme
2 marks for invention activities on Uncertainty in the Mean, and Making Comparisons (something written in your lab book about what you have learned)
1 mark for first plan for measurements
3 marks for pendulum measurements at 10 and 20 degrees, and comparisons
1 mark for plan to improve measurements
3 marks for final high quality measurements and comparisons

Period of a pendulum

Instructions: Measure the period of the pendulum at 10 and 20 degrees and compare.

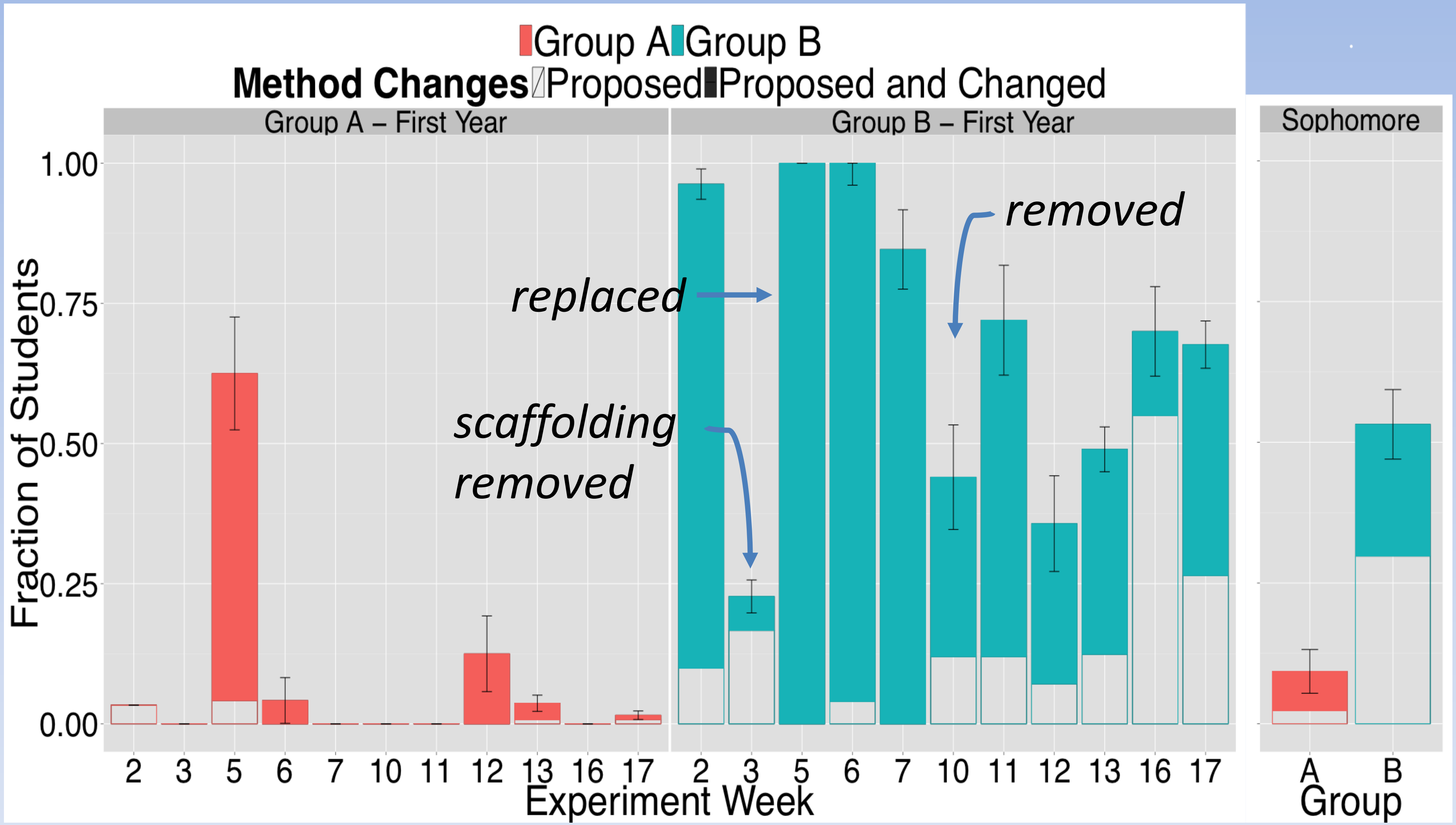
	Students' design	10°	20°	t'-score
Trial 1:	Measure single period 10 times	1.831 ± 0.08	1.805 ± 0.08	0.113
Trial 2:	Measure 10 periods 5 times	1.823 ± 0.008	1.8496 ± 0.008	2.351
Trial 3:	Measure 20 periods 5 times	1.8303 ± 0.004	1.851 ± 0.004	3.659

Faded scaffolding



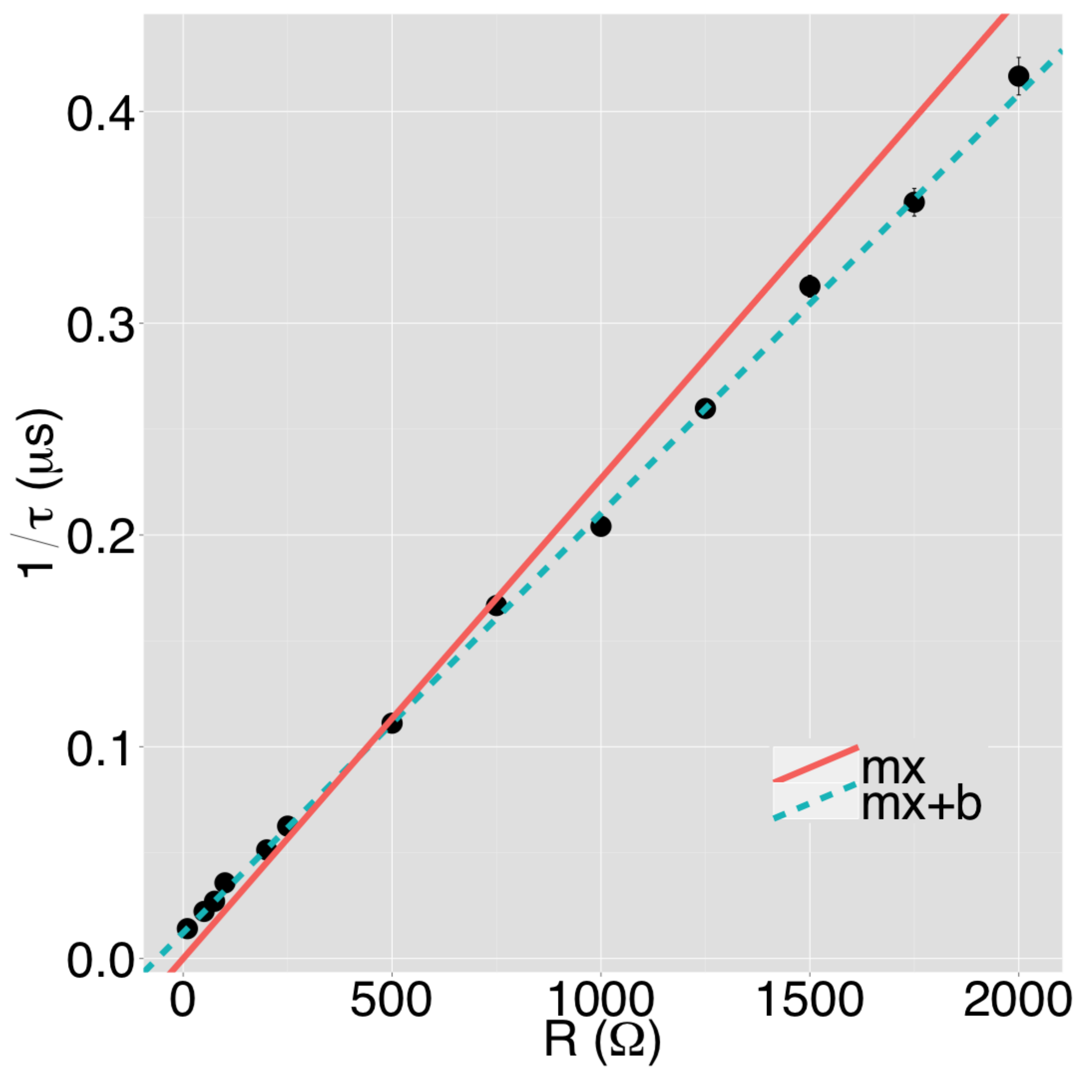
Student support involved instructions and/or grading scheme (so, scale of 0-2 for support of comparing, iterating, and reflecting)

Making improvements becomes a habit



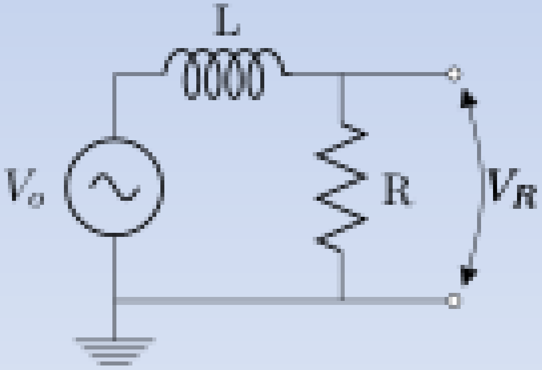
Several weeks of reinforcement needed to achieve sustained improvement – and transfer to second year!

LR Circuits – Lab 17



$$\tau = \frac{L}{R}$$

$$\Rightarrow {}^{1}\!/_{\tau} = \frac{R}{L}$$



Quality of students' reflection on comparisons

Comments in students' notebooks were rated using an adaptation of Bloom's taxonomy.

Level 1 comments remarked on the outcomes of analysis (application without interpretation)

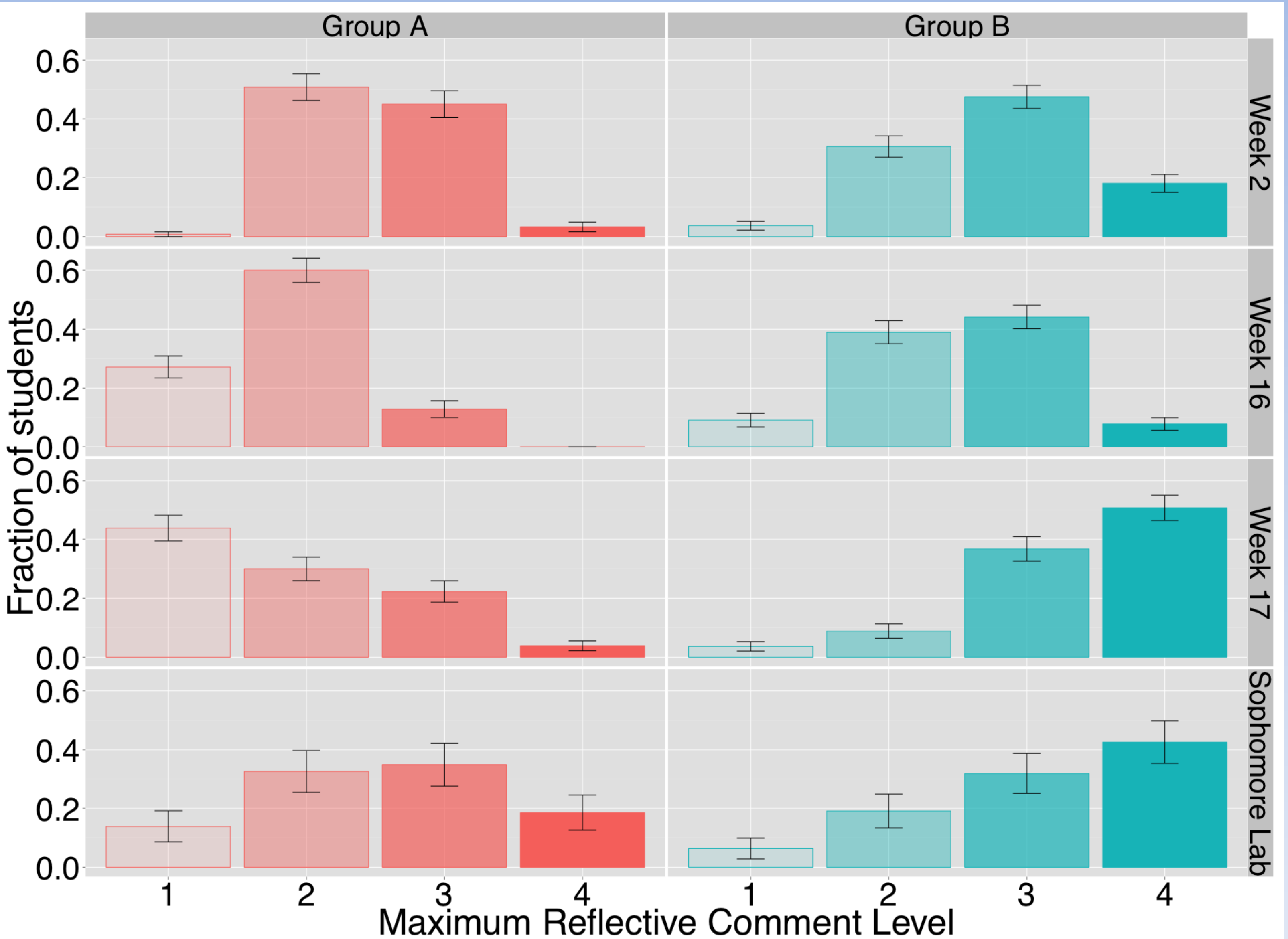
Level 2 comments analyze or interpret data

Level 3 involves synthesis of multiple ideas

Level 4 involves evaluation of the analysis in light of the synthesis

Highest level reached was recorded for each student.

Reflecting on data and results in 4 labs



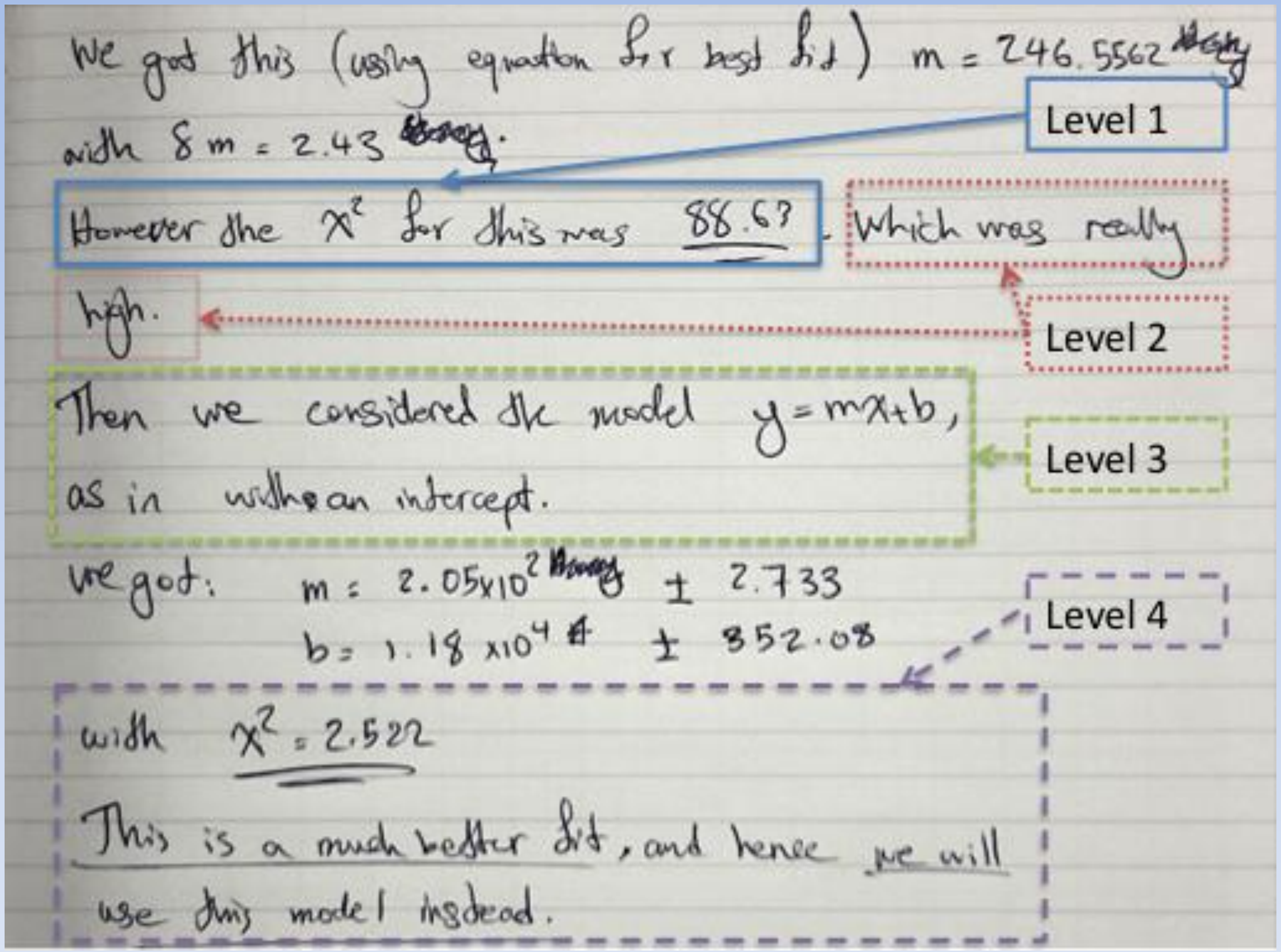
scaffolded model fail (pendulum)

unscaffolded model succeed

unscaffolded model fail (LR circuit)

unscaffolded model fail (LRC circuit)

Coding reflection comments



Core features of SQILabs

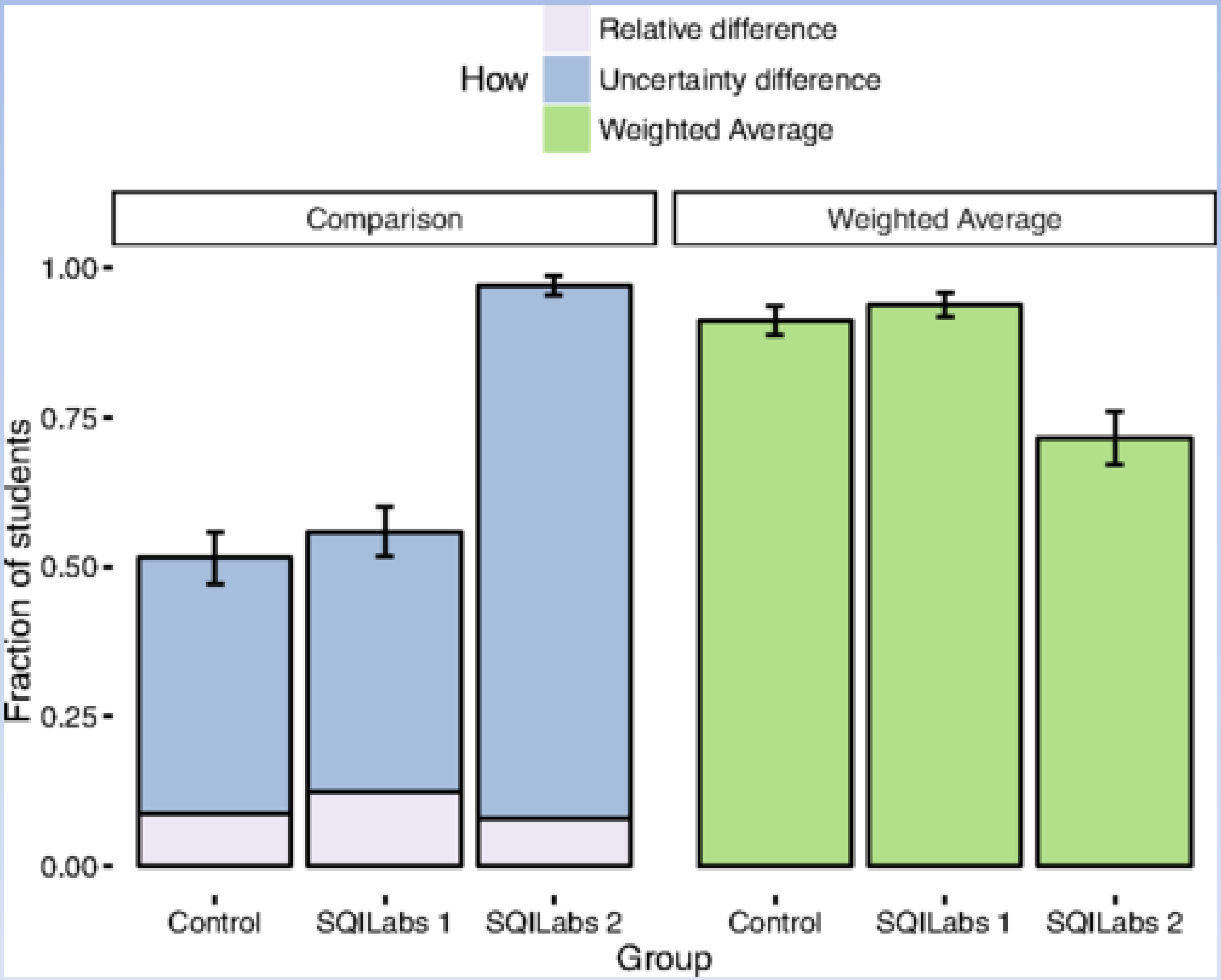
Give the students enough time.

- Relatively simple experiments
- Short enough that they can be refined and repeated
- Experiments must be able to produce high-quality results

Give students free-agency within some constraints

- Comparisons are never confirmatory
- Some comparisons involve a model or assumption that fails
- Both scaffolding and teaching new tools is faded over time
- ***Care with cuing students' frame and task orientation***

Effect of Cuing



Holmes, Kumar, Bonn, Phys. Rev. PER 13, 010116 (2017).

Index of refraction experiment

First SQIlabs version, few students made comparisons

(they had just learned weighted averaging)

Second version, nearly everyone made comparisons (no scaffolding, no new tools)

Conclusions

Give students an environment in which they can do authentic scientific inquiry, but constrained and supported in ways that keep it productive.

Support is sustained in order to develop scientific habits (making quantitative comparisons and iterating/improving)

Support can be faded over time, leaving lasting improvements.

Students eventually take ownership of their own learning in the laboratory, with striking gains in their scientific reasoning.

Holmes, Wieman, Bonn, PNAS 112, 11199 (2015).

Extra slides for questions

SQILab Design Principles

Learn new data tools at a pace that allows practice and synthesis

Experiments must be able to produce high-quality results

Experiments are simple and short enough to do more than once

Comparisons are rarely confirmatory

Some comparisons involve a model or assumption that fails

Support expert-like behaviours with explicit scaffolding

Careful alignment of grading and goals

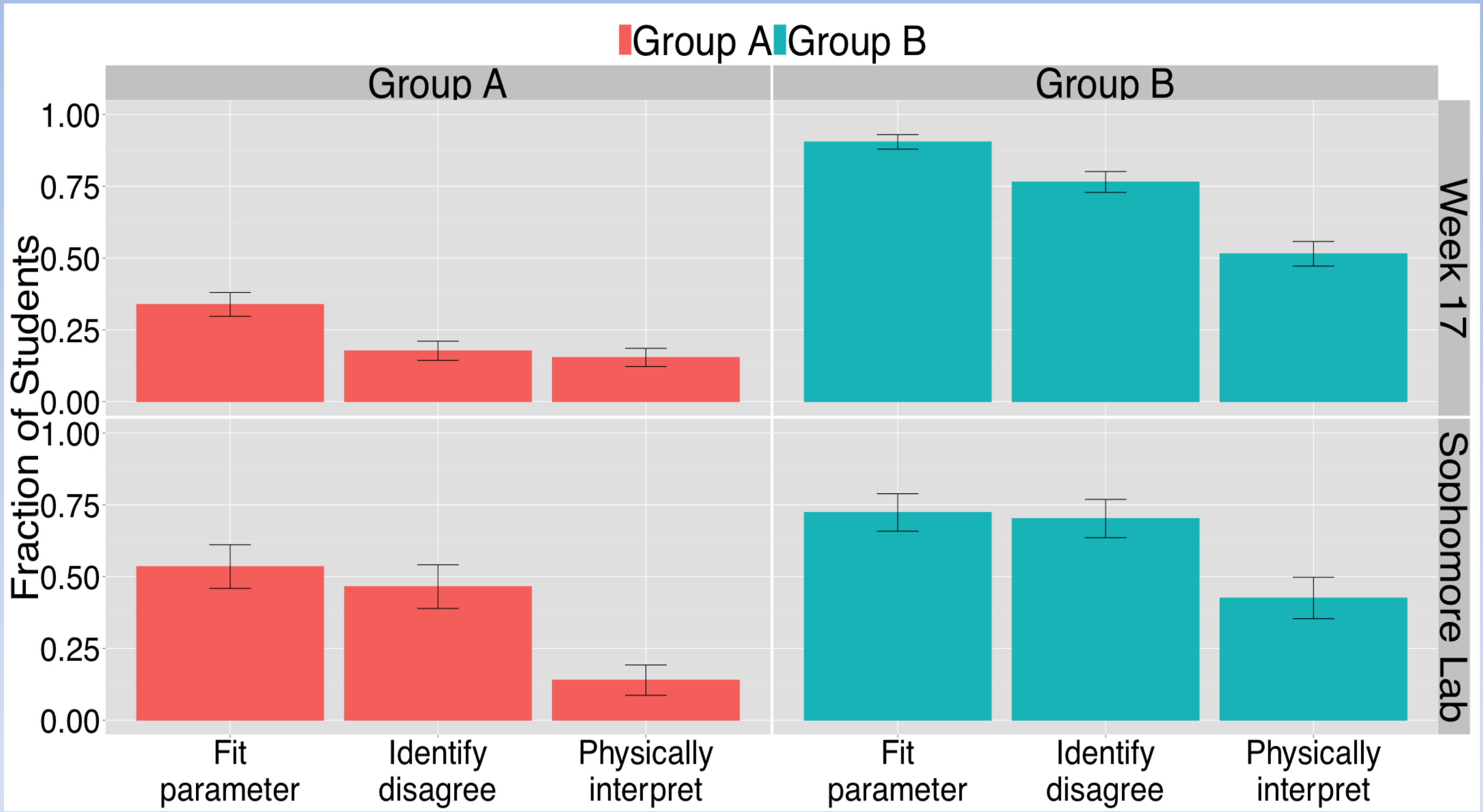
Fade scaffolding over time

Near the end, practice without learning new tools

"When I'm reading about something or solving physics problems or just reading about physics concepts, the idea of me being a physicist in that sense is very far fetched…[the lab] helped me think about a bunch of data that I have in front of me, that looks like chaos, in a more scientific way…

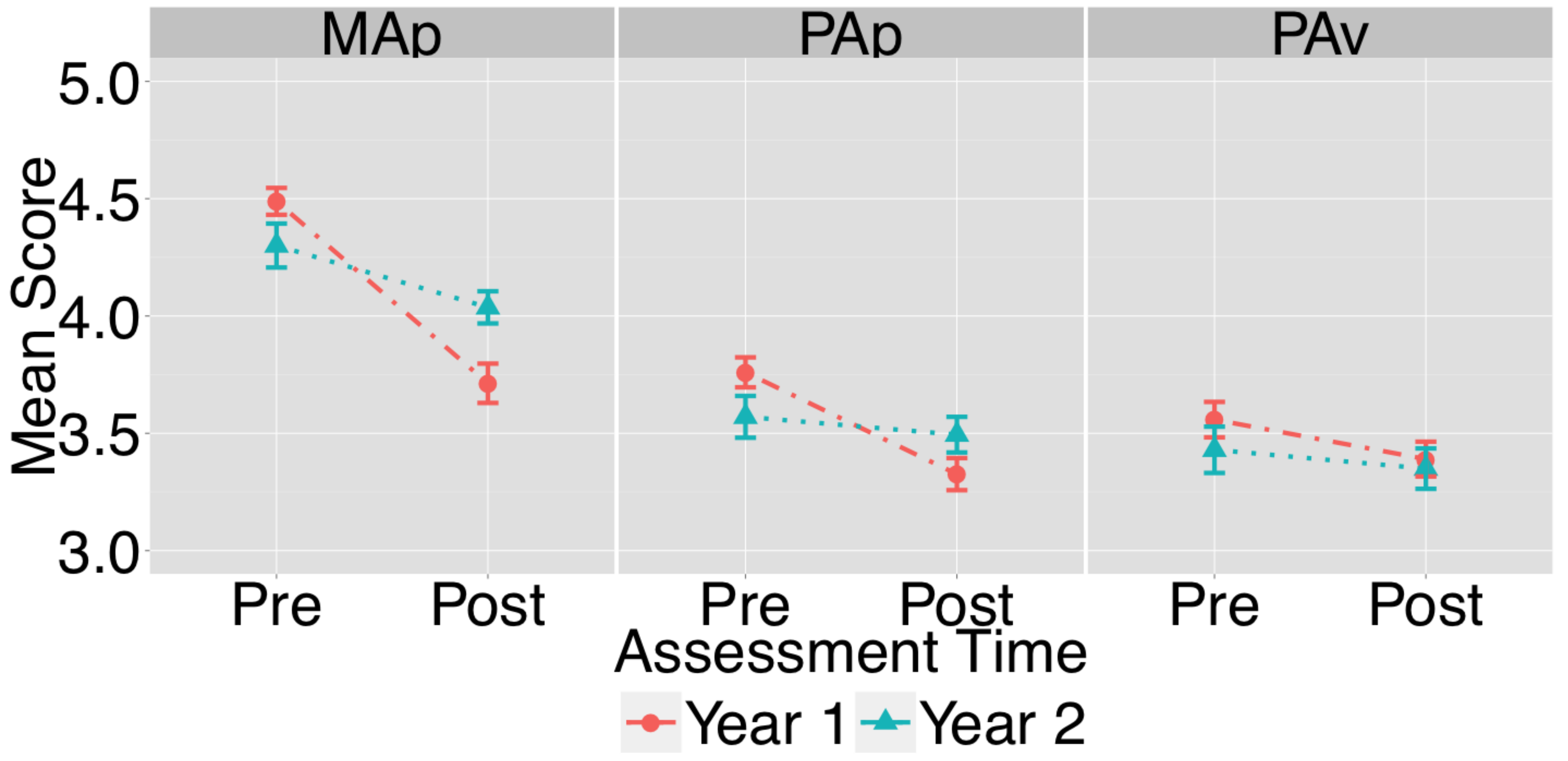
[The lab] integrates everything so much more and it helps me see myself as a scientist way more than all my other classes, because those are just putting information… giving me information, rather.. It helps me actually reach in and realize, 'oh, this makes sense! I can actually do this too,' rather than just memorize a textbook."

Evaluating transfers to sophomore year

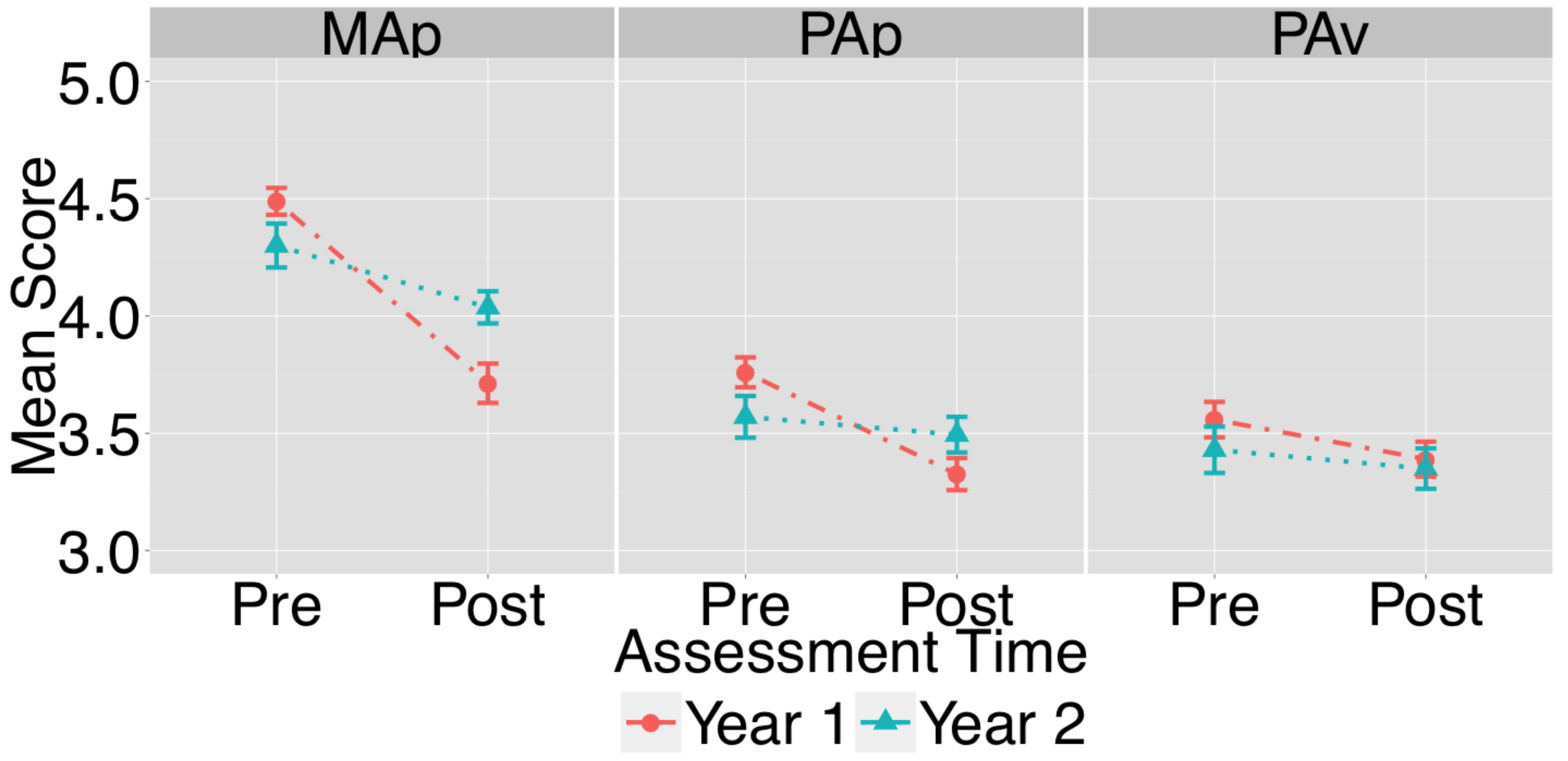


Evaluation and physical interpretation of a model transfers into a different experiment 6 months later

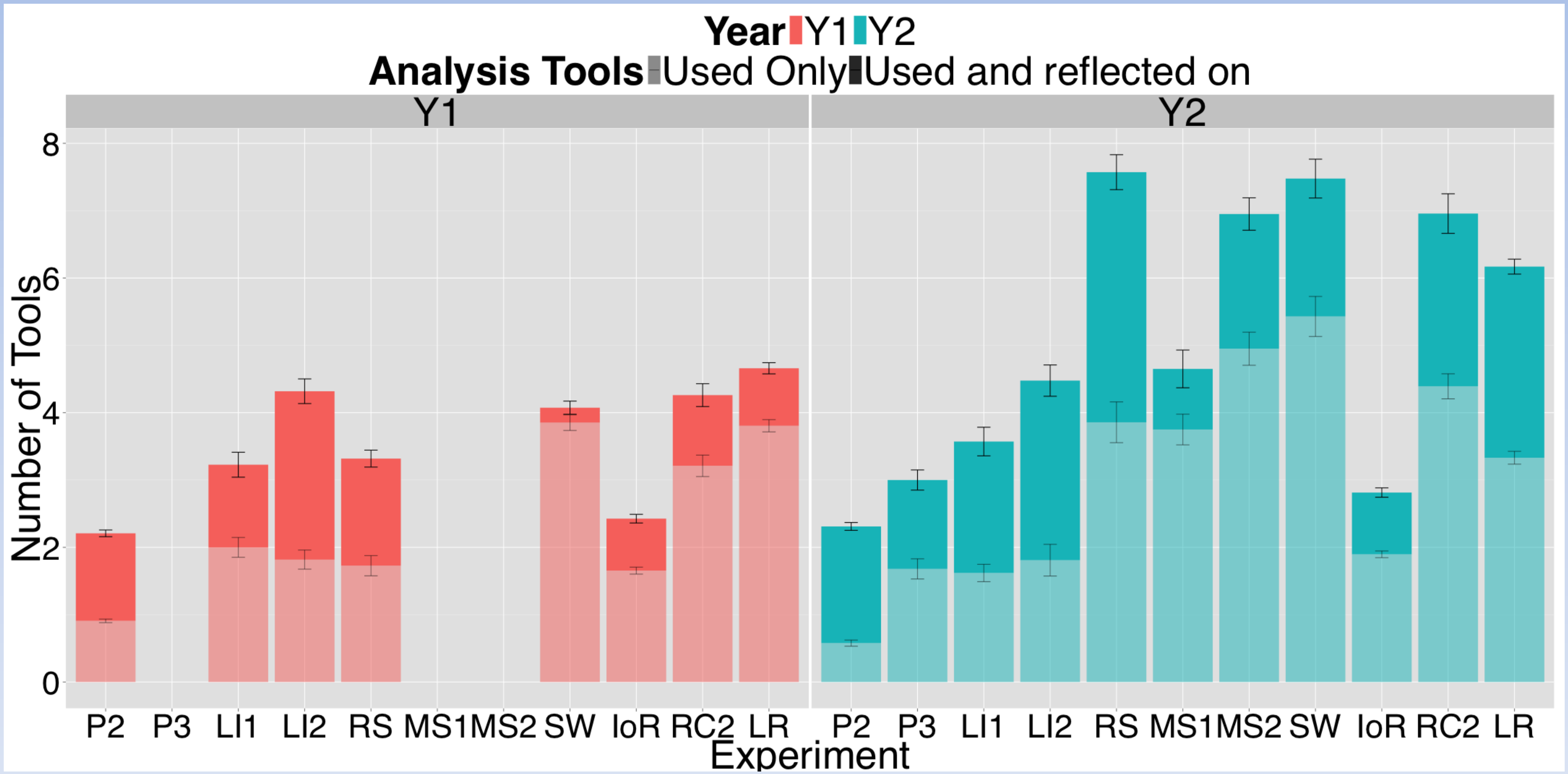
Attitudes and epistemologies

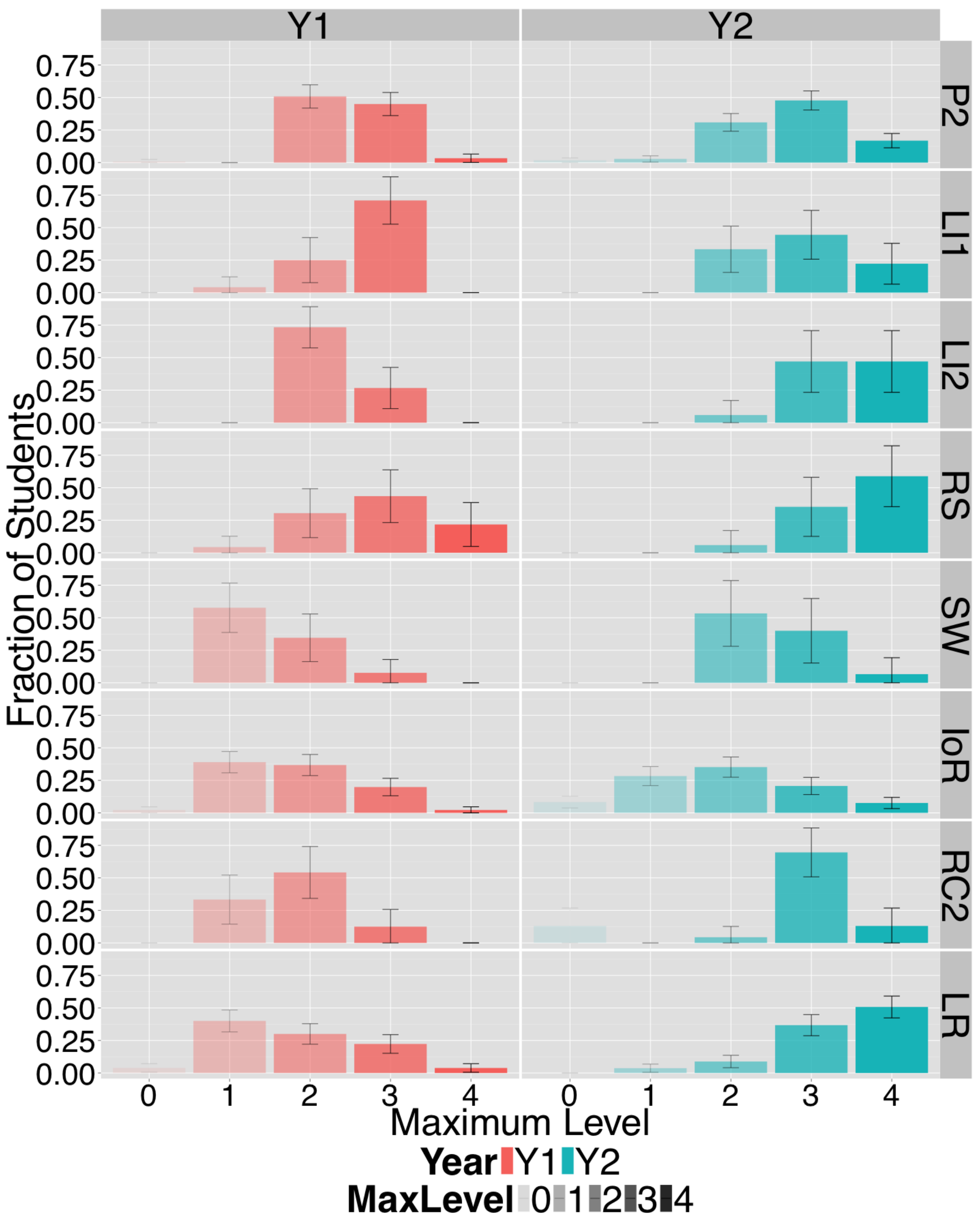


Attitudes and epistemologies

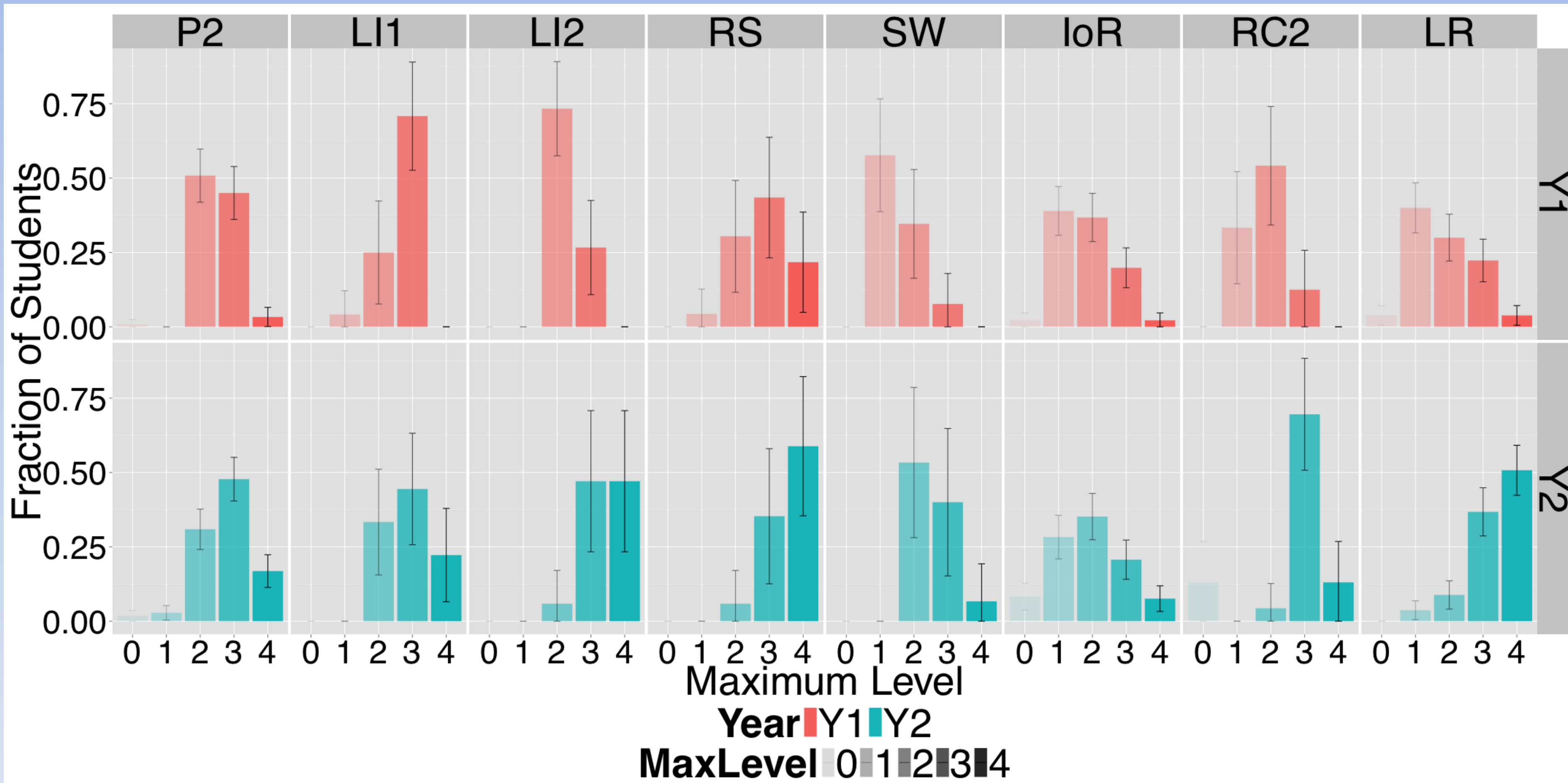


Reflecting on data and results





Y1
Y2
P2
LI1
LI2
RS
SW
IoR
RC2
LR
Fraction of Students
0.00
0.25
0.50
0.75
0
1
2
3
4
Maximum Level
Year Y1 Y2
MaxLevel 0 1 2 3 4

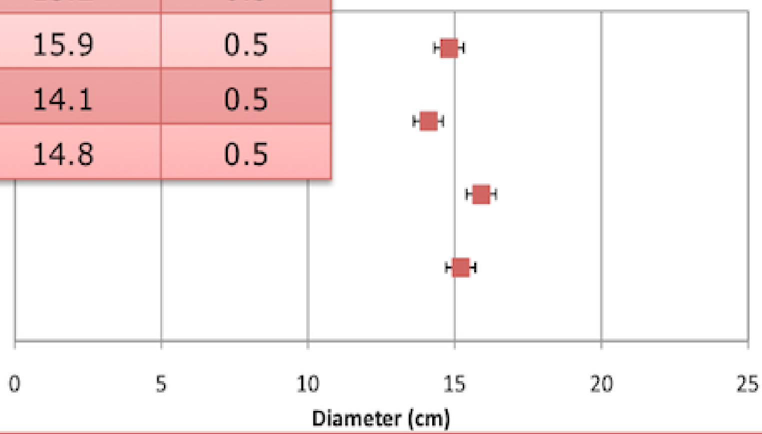


P2
LI1
LI2
RS
SW
IoR
RC2
LR
Y1
Y2
Fraction of Students
0.00
0.25
0.50
0.75
0 1 2 3 4
Maximum Level
Year Y1 Y2
MaxLevel 0 1 2 3 4

Ostrich eggs invention

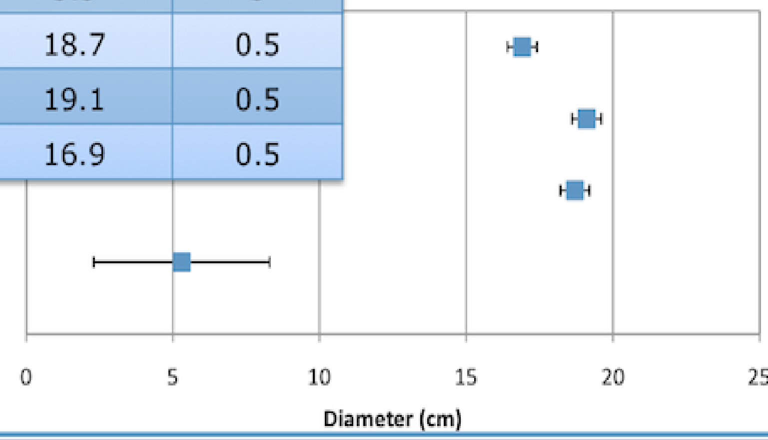
Chef A

Diameter (cm)	$\delta_{Diameter}$(cm)
15.2	0.5
15.9	0.5
14.1	0.5
14.8	0.5



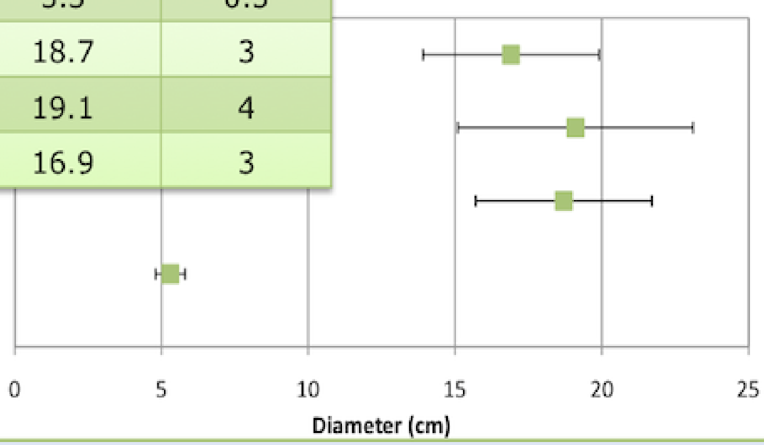
Chef B

Diameter (cm)	$\delta_{Diameter}$(cm)
5.3	3
18.7	0.5
19.1	0.5
16.9	0.5



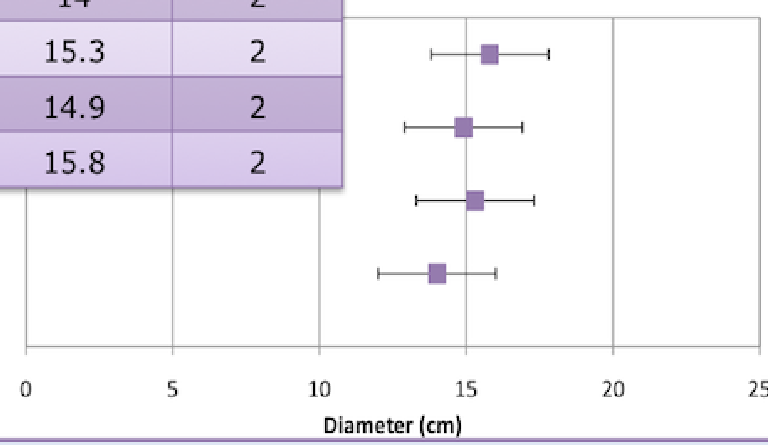
Chef C

Diameter (cm)	$\delta_{Diameter}$(cm)
5.3	0.5
18.7	3
19.1	4
16.9	3



Chef D

Diameter (cm)	$\delta_{Diameter}$(cm)
14	2
15.3	2
14.9	2
15.8	2



Hubble Constant

- Riess et al. used revised classical distance ladder techniques to find:

$$H_0 = 73.8 \pm 2.4 \text{ km/s/Mpc}$$

- Compare to value derived from ΛCDM model fit to CMB + BAO data:

$$H_0 = 68.76 \pm 0.84 \text{ km/s/Mpc}$$

Do these two measurements agree or disagree?

Adapted from Gary Hinshaw – PHAS Colloquium 2012

Hubble Constant

- Riess et al. used revised classical distance ladder techniques to find:

$$H_0 = 73.8 \pm 2.4 \text{ km/s/Mpc}$$

- Compare to value derived from ΛCDM model fit to CMB + BAO data:

$$H_0 = 68.76 \pm 0.84 \text{ km/s/Mpc}$$

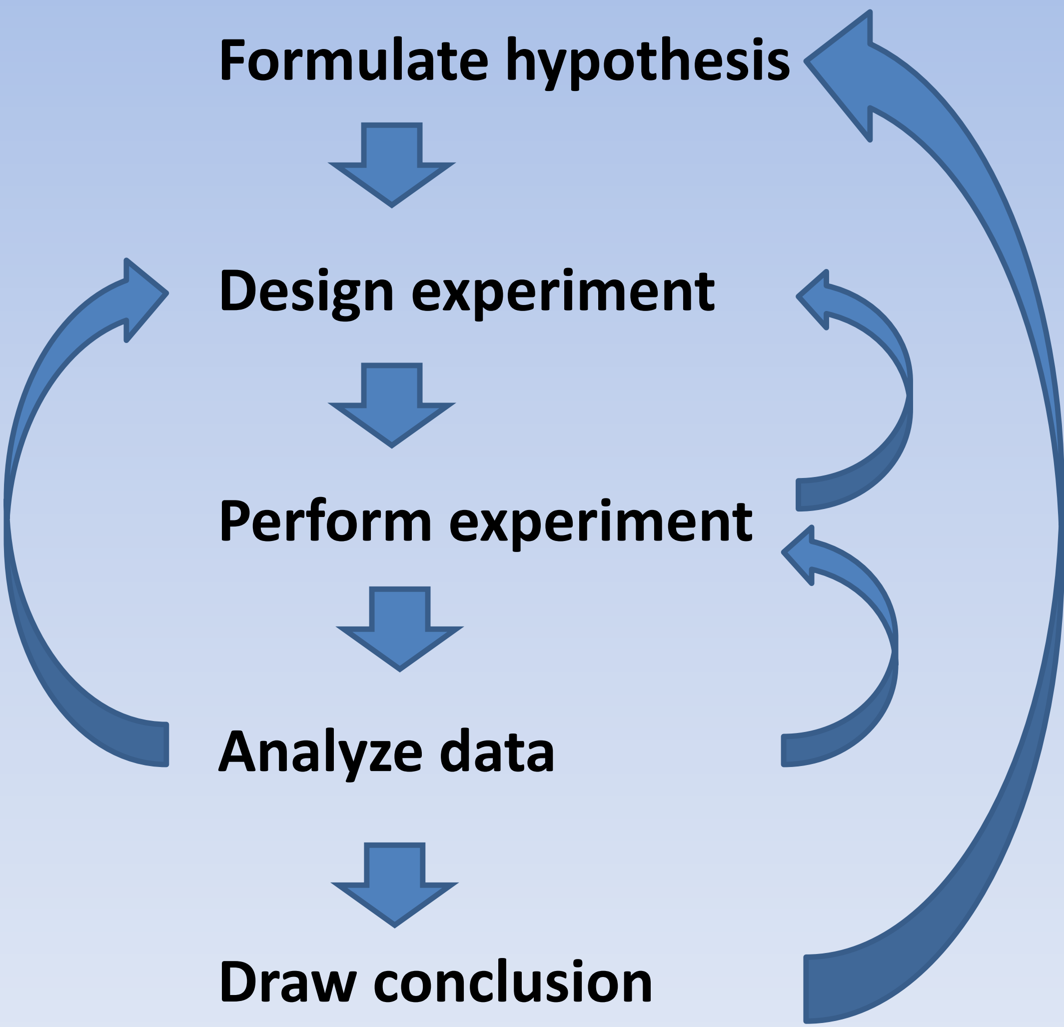
- Mild tension, but consistent at the 2-sigma level. Justified to include independent H_0 data into full fits.

Adapted from Gary Hinshaw – PHAS Colloquium 2012

Technical tools

Estimation of uncertainty (random, analog, digital)
Histograms
Mean, standard deviation, standard uncertainty of the mean
Propagation of uncertainty
Comparing measured values with a t-score
Weighted mean
Scatter plots
Semi-log plots and exponentials
Log-log plots and power laws
Building a model (function) and comparing to data (numbers)
Residuals
Linear least-squares fitting using weighted chi-squared
Estimation of uncertainty in fit parameters

Obstacle #2
The ‘scientific method’

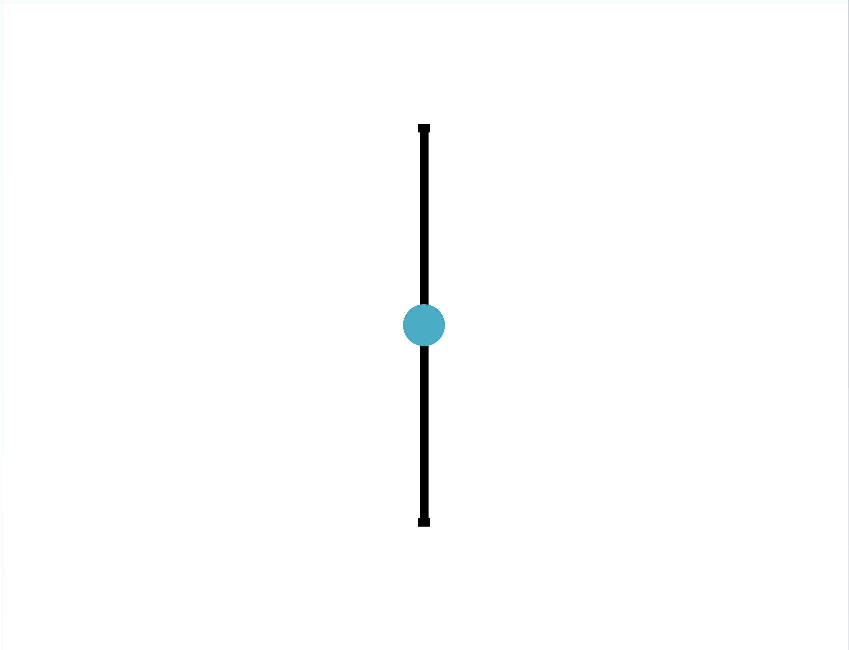


Students do not typically appreciate the iterative nature of this process.

Obstacle #1
Measurement uncertainty

We can make the students' misunderstanding even worse:

- Using the word 'error', when we mean 'uncertainty'
- Giving them experiments where they regularly fail
- Doing only confirmatory experiments



- $A \pm \delta_A$ -> $[A+\delta_{A,}\ A-\delta_A]$

  (hard limits)
- $A \pm \delta_A$ -> $A+\delta_A$ or $A-\delta_A$

  (binomial theorem)

Obstacle #1
Measurement uncertainty

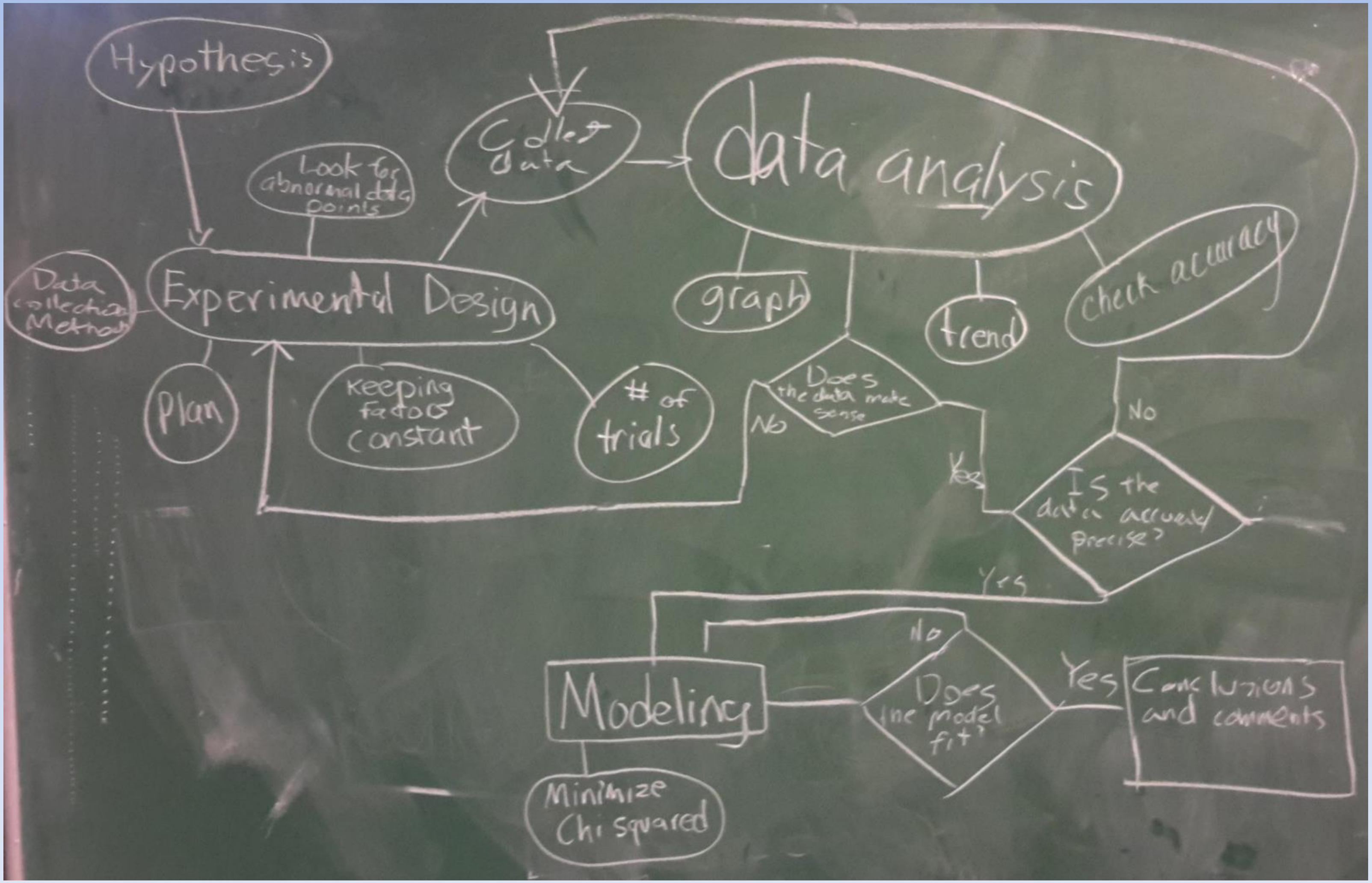
Students come into first year with a problematic understanding of the word error.

$$\%error = \frac{Measured\,value - True\,value}{True\,value} \times 100\%$$

Where
'True value' means precise result of an expert scientist
'Measured value' means lousy measurement of a novice student

Bad for students' understanding of measurement, bad for their beliefs about the nature of science

Half-way through the course, they know about iterative loops



Structured Quantitative Inquiry: SQILabs

Move away from labs as a support to learning physics concepts.

FIRST ERA

- ***Develop a functional understanding of measurement uncertainty***
- ***Learn a set of broadly applicable data-handling skills***

SECOND ERA - A RELATED SET OF METACOGNIVE GOALS

- ***Develop expert-like habits of mind and scientific reasoning***
  - Meaningful reflection on the quality of data
  - Meaningful reflection on fit between data and model
  - Use the iterative nature of science
  - Develop confidence that they can do high-quality measurements