

Doing Science Or Doing A Lab? Engaging Students With Scientific Reasoning During Physics Lab Experiments

N.G. Holmes and D.A. Bonn

Dept. of Physics & Astronomy, University of British Columbia, 6224 Agricultural Rd, Vancouver, BC, V6T 1Z1

Abstract: Undergraduate teaching labs often include expectations for students to use various sensemaking and reasoning behaviours when conducting an experiment. These expectations, however, are often unsupported in the explicit lab structure, and, as such, students develop distinctions between behaviours associated with “doing a lab” or “doing science.” This study examines if and how students engage in reflection and evaluation of results during an experiment that involves two sources of systematic error. While many students reflected on their results and identified the source of the larger systematic error, very few did so for the smaller one. In fact, for the latter case, many students reported significantly inflated uncertainties, effectively hiding the systematic error altogether. We use these results, in-class observations, and post-lab interviews to describe how and why students failed to demonstrate authentic scientific inquiry behaviours during the lab.

Keywords: laboratory, experimental physics, systematic errors, scientific reasoning.
PACS: 01.40.Fk, 01.40.G-, 01.40.gb

INTRODUCTION

When defining the goals of the introductory physics laboratory, the American Association of Physics Teachers placed emphasis on understanding and developing skills in scientific measurement and experimentation [1]. Students, however, often define scientific measurement and the nature of science differently in a research setting compared to a classroom context [2]. For example, it was reported that only when prompted to reflect would students recognize that their answers to in-class problems were physically unrealistic [3]. This was attributed to students distinguishing two epistemological frames: calculating values for the sake of completing a physics problem versus calculating values that represent real quantities. In a first-year laboratory, these two frames might manifest themselves in a frame of doing a measurement for the sake of getting a lab done versus actually doing a measurement of a physical quantity. Furthermore, it has been shown that students consider their measured results separate and distinct from those achievable by scientists (for example, compared to *true* or *accepted* values) [4].

Many researchers have begun to explore these sorts of apparently conflicting frames through surveys or interviews with students outside of classroom activities. These methods provide a decontextualized image of students’ beliefs about the nature of science and measurement [2,4,5]. They are limited in that they do not probe whether or how students carry these

frames with them into activities such as labs, problem solving, or tutorials.

Here we explore these issues of framing and measurement perceptions in a first-year lab course at the University of British Columbia. In a particular experiment, students measured a single physical quantity using three different methods, two of which were prone to systematic errors that almost all students would encounter. The apparatus had sufficient precision that the systematic errors were identifiable and could be resolved by an expert. The present study examines if and how students use scientific reasoning to correct the systematic errors. We use in-class observations and interviews outside of class to suggest reasons for their actions. These in turn motivate instructional approaches to address how to improve students’ expertise in addressing conflicting results.

THE LAB: INDEX OF REFRACTION

This study involved 136 students in a first-year physics lab course that runs for two semesters. The learning goals in the course focus primarily on developing skills related to measurement, uncertainty, graphing, and statistical analysis [6]. Halfway through the second semester of the course, students worked in pairs or groups of three to determine the index of refraction (n) of a plexiglass prism through the application of Snell’s Law, the critical angle for Total Internal Reflection, and Brewster’s Angle. The students had received instruction earlier in the year on assessing whether measurements were in agreement

and on quantifying uncertainties from digital and analog scales and due to statistical fluctuations [7].

Snell's Law

For the first measurement, students were asked to determine n of the prism through Snell's Law (SL):

$$n = \frac{\sin(\theta_{\text{incident}})}{\sin(\theta_{\text{refracted}})} \quad (1)$$

(where $n_{\text{air}}=1$). They were asked to use an incident angle of 60° , measure the angle of the refracted beam, and then determine n and an estimate of its uncertainty from Equation (1). The approximate orientation for the apparatus and relative positions of the beams were as in Fig. 1. The incoming beam entered the flat side of the prism along the 0° -line of the protractor. For an incident angle of 60° the position of the normal line is rotated to align with the 60° -line of the protractor. As such, the refracted beam exited the plexiglass along the $24.3^\circ \pm 0.5^\circ$ mark of the protractor, setting the refracted angle at $35.7^\circ \pm 0.5^\circ$ (relative to the normal line). Using Equation (1), n would be 1.48 ± 0.02 . The fixed nature of the incident beam (rather than the normal line) along the 0° -line of the protractor led to a common systematic error that measured the refracted angle as 24° , giving n around 2.13.

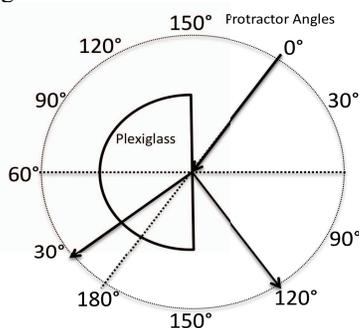


FIGURE 1. The diagram represents a schematic of the plexiglass and protractor orientation for the SL measurement. Approximate beam angles use two 0° to 180° scales on a fixed protractor.

Total Internal Reflection

For the second measurement, students used the properties of Total Internal Reflection (TIR) to determine n . They were asked to measure the critical angle of incidence, θ_{critical} , beyond which the incident beam is only reflected, with no refracted beam, and then determine n and an estimate of its uncertainty from:

$$n = \frac{1}{\sin(\theta_{\text{critical}})} \quad (2)$$

In the SL measurement, the refracted beam is approximately 1 mm wide. As the incident angle approaches the critical angle, however, the refracted beam spreads to nearly 10 cm in width. With appropriate identification of θ_{critical} as the point at which the centre of the beam is about to disappear, approximately $42.5^\circ \pm 0.5^\circ$, Equation (2) gives n of 1.48 ± 0.01 . Instead, a common systematic error is to take θ_{critical} as the point where the refracted beam has entirely disappeared, around 45° , giving n of 1.41.

Brewster's Angle

For the third measurement, students used a polarizer to determine n and its associated uncertainty using Brewster's angle (BA), the angle of incidence at which the reflected beam is completely polarized:

$$n = \tan(\theta_{\text{Brewster's}}) \quad (3)$$

This measurement did not include any common systematic errors and, in general, students were able to accurately measure $\theta_{\text{Brewster's}}$ at $56^\circ \pm 1^\circ$, with Equation (3) giving n of 1.48 ± 0.06 .

RESULTS

Students' values for n from the SL and TIR measurements were extracted from their lab books and each categorized by researchers as being accurate (within 1.48 ± 0.03), inaccurate (outside this range), or accurate after making a change to the calculation or measurement¹. Original and final values through the changes made in the latter case were discernible in students' lab books since students simply crossed out the original values and replaced them with final ones. This suggests students tried to hide their mistakes rather than provide a clear and distinct description of how they managed to correct their errors.

In the following paragraphs, these quantitative results present aspects of *what* students did in the lab and explanations of *why* are connected to student interviews and in-class observations.

For SL, students' original and, where appropriate, corrected values were distributed bimodally as in Fig. 2 with 49% of students making the identified systematic error (demonstrated as the second peak around $n=2$ for the initial values) and 58% of those students correcting the error (shown through the larger peak near $n=1.48$ for the final values). Several strategies for making corrections were observed in the lab or discussed during interviews. The most common strategy, also implicitly supported in the lab structure, was by comparing the three measurements. Students identified

¹ BA measurements were also extracted but not categorized since there was no systematic error.

the SL error since it produced a value that differed from the other two by nearly 50%. Students also conferred and compared values with other groups or used the provided formulas to make predictions about subsequent measurements. Interviews suggested, surprisingly, that many students thought that these tactics (talking to other groups or comparing calculated predictions to measurement) were forms of cheating, which may have limited whether they could effectively reflect on their results. Other reflection techniques used commonly in science were not observed, such as to keep precision as high as possible to reveal inaccuracies or to attempt measurements in multiple ways (such as to use a different incident angle for SL).

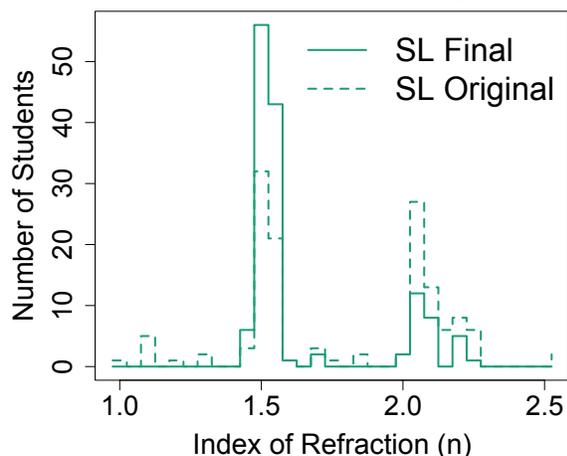


FIGURE 2. The graph shows the distribution of students' original (dashed line) and final (solid line) n values from the Snell's Law measurement.

The remaining students (20% of the class) did not effectively reflect on their results and left the lab with the large SL systematic error. While a plausible explanation is that students did not have enough time to reflect or re-measure, over 75% of students had left the lab with 45 minutes remaining in the lab period and no students stayed for the full three hours. Many of the previous labs, however, were short on time and required students to work methodically and efficiently without reflection. As one student described, she often had to “stop thinking and write something down,” a habit that may have carried through to this lab. In addition, the shorter lab meant that students would weigh the relative value in spending more time to reexamine their results over leaving the lab early.

60% of students made the TIR error and only 2% of them corrected it, compared to 58% who corrected the SL error. Interviews suggested that this difference was due to the relative sizes of the errors: the SL measurement differed from the BA measurement by 50%, whereas the TIR error differed by only 5%. Comparing data for the three measurements from all students (Fig. 3) presents tension between the TIR data

with other measurements, as the TIR peak is centered slightly left of the SL and BA peaks. For an individual student without the full dataset, however, any disagreement associated with their TIR measurement may have seemed unimportant relative to the size of the SL error. In addition, many students increased their uncertainties in the TIR measurement to account for the spreading of the refracted beam as it approached the critical angle, which then meant that the values agreed, though within inflated uncertainty ranges.

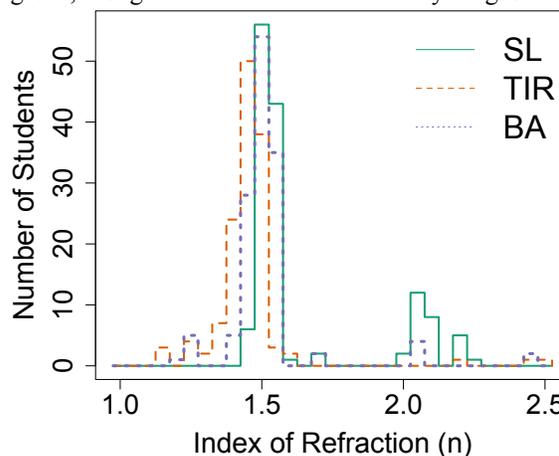


FIGURE 3. The graph shows the distribution of students' final n values for each of the three measurements in the index of refraction lab.

The course instructor and teaching assistants (TAs) performed the experiment and their measurements can be used as a benchmark for evaluating the precision of the students' measurements. In general, students reported uncertainties that were two to four times larger than the uncertainty obtained by the TAs and instructor (through one-sample t -tests): SL, $t(134)=5.68$, $p<.001$; TIR, $t(134)=3.66$, $p<.001$. Furthermore, students who completed the lab with inaccurate measurements reported uncertainties that were around three times larger than students who completed the lab with accurate measurements (initially or corrected): SL, $t(35)=3.70$, $p=.001$; TIR, $t(80)=2.21$, $p=.030$. This supports previous evidence that students merge the concepts of accuracy and precision [4] and additionally suggests that students report poor precision to reflect inaccuracies. In one interview, a student, who had assigned plus or minus 3° uncertainty to each of his angle measurements, justified this in that he and his partner “didn't try to be accurate, [they] tried to be safe.” This reasoning suggests students perceive greater worth to having the values agree than to measuring accurately and precisely, which is contrary to scientific practice.

Research also suggests that many students believe that they, as students, cannot measure accurately or precisely. In fact, it has been shown that students

attribute uncertainty only to their own measurements, whereas so-called accepted values, presumably those determined by scientists, have no associated uncertainty [5]. This seems to have been transferred to this lab in that students expected to make poor quality measurements, leaving no need to reflect on their results to try to improve their values. One student, who corrected their SL measurement, remarked at being “surprised by how close the values were,” especially since he had previously described the plexiglass as simply “a lump of acrylic.” Other students mentioned not trusting equipment or their own data collection.

Earlier in the course, students had received instruction on how to assess whether values with uncertainties agreed, for example through range overlap. This sort of procedural instruction can often cause students to follow these algorithms systematically without sufficient sensemaking [8]. In our case, students drew conclusions that their values did not agree without reflecting on what it meant for nominally equivalent values to disagree.

DISCUSSION

This study examined how students deal with systematic inaccuracies and disagreements between nominally equivalent measurements and how these relate to their ideas about the nature of science. Several themes emerged regarding why students may fail to use scientific reasoning and reflection strategies in such situations in an introductory physics lab.

Firstly, students engaged in inauthentic scientific reasoning such as quoting inappropriately large uncertainties to obtain agreement between measurements. This may be attributed to a disconnection between the nature of science in a research lab and in a teaching lab. Students applied the previously reported assumption that there is inherent “human error” in their measurements, compared to ideal measurements by scientists [5]. They assumed that their values did not agree due to contextual limitations (they as students in a lab) as opposed to correctable systematic errors. This assumption may stem from a lack of experiences in making high quality measurements, which causes them to believe that they, as students, cannot measure well or precisely.

Secondly, the lab structure itself may not have supported authentic scientific reasoning. In particular, since assessment in the lab placed higher emphasis on products over process (often, marks were rewarded for the presence of the measurements and conclusions in their report, rather than for their quality), students saw higher value in getting “safe” values down quickly over spending time taking better data. In this way, reflecting on and evaluating the quality of their results was not necessary for success. Furthermore, the lab

activities did not explicitly teach students about useful tactics for reflecting and evaluating their work. Students even believed that some tactics were dishonest even though they are regularly used in scientific research. Without these strategies, it is understandable that they were not able to use appropriate reasoning behaviors.

Future work is targeted around restructuring the course to address the issues above. This will include providing multiple opportunities and ample time to reflect on results and improve the quality of measurements, explicit support in developing effective reasoning strategies, and new assessments that reward students for engaging in scientific behaviors. The goal is for scientific reasoning to become a valuable habit in the lab and for students to be empowered to do a good job, to understand that they can do a good job, and to repeat measurements to increasingly do a better job.

When we revisit the index of refraction lab we will observe whether new structures in previous labs will have affected the outcomes described here. We anticipate that engaging students in more authentic scientific experiences will start to break down their distinctions between science and the lab.

ACKNOWLEDGMENTS

We would like to thank James Day and Ido Roll for their support and assistance during the development of this study, and Andrew Macdonald and Amy Robertson for the valuable feedback on the manuscript. This research was supported by UBC’s Carl Wieman Science Education Initiative.

REFERENCES

1. American Association of Physics Teachers, *American Journal of Physics* **66**, 483–485 (1998).
2. M.-G. Séré, M. Fernandez-Gonzalez, J.A. Gallegos, F. Gonzalez-Garcia, E.D. Manuel, F.J. Perales, and J. Leach, *Research in Science Education* **31**, 499–523 (2001).
3. M.A. Martinuk, Ph.D. dissertation, University of British Columbia, Vancouver, 2012.
4. S. Allie, A. Buffler, B. Campbell, and F. Lubben, *American Journal of Physics* **20**, 447–459 (1998).
5. T.S. Volkwyn, S. Allie, A. Buffler, and F. Lubben, *Physical Review Special Topics - Physics Education Research* **4**, 010108 (2008).
6. J. Day, and D.A. Bonn, *Physical Review Special Topics - Physics Education Research* **7**, 010114 (2011).
7. S. Allie, A. Buffler, B. Campbell, F. Lubben, D. Evangelinos, D. Psillos, and O. Valassiades, *The Physics Teacher* **41**, 394 (2003).
8. B.M. Zwickl, N. Finkelstein, and H.J. Lewandowski, in *Physics Education Research Conference, AIP Conference Proceedings*, **1513**, 442–445 (2012)