Claims, evidence and reasoning in the introductory mechanics lab

Andrew Pawl
Engineering Physics Department, University of Wisconsin-Platteville, Platteville, WI, 53818

Introductory mechanics is classified as a general education laboratory science at many colleges and universities. General education outcomes often include the ability to reason from evidence or justify claims with evidence. These skills are also central components of the Next Generation Science Standards for K-12 education. In contrast to this mandate to focus on evidence-based reasoning, both the teaching and the assessing of the ability to reason from evidence are often implicit rather than explicit parts of the introductory mechanics laboratory curriculum. This article reports the first results of an ongoing attempt to scaffold the learning of reasoning from evidence and to make the assessment of this skill explicit by employing the "Claim, Evidence and Reasoning" framework in a college-level introductory mechanics laboratory.

2022 PERC Proceedings edited by Frank, Jones, and Ryan; Peer-reviewed, doi.org/10.1119/perc.2022.pr.Pawl

I. INTRODUCTION

Introductory mechanics is classified as a general education course at many colleges and universities. At the small (about 6500 students) state university where I teach, the course is designated as a Natural Science general education course. As part of maintaining this designation my department is required to perform assessment of the university's Critical Thinking student learning outcome. The department is specifically required to choose two separate aspects of the Critical Thinking outcome to assess and we have selected:

1. Evidence, defined as the "ability to select and use information to investigate a point of view, position, or hypothesis".
2. Conclusions, defined as the "ability to articulate consequences and implications of a point of view, position, or hypothesis".

The idea that one important reason for studying science is to learn how to support conclusions with evidence is also reflected in curricular standards for K-12 education. The Next Generation Science Standards for K-12 education incorporates eight "practices" which are central to the implementation of the standards. Practice #7 is "Engaging in argument from evidence" [1]. These eight practices were adopted verbatim from the National Research Council's (NRC) report entitled "A Framework for K-12 Science Education" [2]. In justifying the inclusion of practice #7, the NRC's report states that science relies on a "process of reasoning that requires a scientist to make a justified claim about the world". Further, the report makes the case that this skill is relevant for non-scientists because "the scientist and the citizen alike must make evaluative judgments about the validity of science-related media reports and their implications for people's own lives and society". This rationale supports the inclusion of outcomes related to reasoning from evidence in general education curricula at the university level as well.

There are many ways to interpret the term "evidence", but in the context of a Natural Science general education requirement it is often assumed that experimental data is meant and the requirement can only be fulfilled by taking one or more courses involving a laboratory. In the NRC's report, it is not explicitly stated that the evidence in practice #7 refers to experimental data, but the associated goals for practice #7 include items such as: "Recognize that the major features of scientific arguments are claims, data, and reasons" and: "Construct a scientific argument showing how data support a claim". These goals suggest that the NRC equates the term "evidence" with scientific data in the context of science education.

If students are to be assessed on their ability to use experimental data as evidence to support a claim, the laboratory must be a key component of this assessment. The laboratory, however, is also performing another function. It is used to reinforce conceptual material presented in the lecture portion of the course. If the focus of the laboratory is leading students through a step-by-step verification of a complex theory then the instructor has not required students to argue from evidence. Instead, as described by Holmes and Bonn, students can effectively go through the labs in a "plug-and-chug frame" with "little to no sensemaking" [3]. The lab may demonstrate that students can successfully take data as requested and perform a prescribed mathematical analysis, but it has not answered the question of whether the students can provide the reasoning that links the data gathered to the theory being tested. The instructor has provided that reasoning and assumed the students can follow the argument. This assumption likely invalidates the use of the laboratory as an assessment of a student's ability to justify claims with evidence. More importantly, the lack of an explicit focus on arguing from evidence probably renders the laboratory useless as a means of teaching this skill. Given that the stated purpose of a general education course is to teach the associated student learning outcomes not simply assess them, it is important to explicitly incorporate arguing from evidence into the introductory mechanics lab so students can see that this skill is a component of the class and gain facility through practice.

This is a long-standing issue and several approaches have been developed to address it (see, *e.g.*, [4–10]). One framework designed to explicitly scaffold students as they learn the skill of arguing from evidence is the "Claim, Evidence, Reasoning" (CER) rubric [11, 12].

- A claim is a statement that answers a question or problem.
- Evidence is scientific data that supports the claim.
- Reasoning provides a justification for why or how the evidence supports the claim.

In this work, I describe the implementation of the CER rubric in a lab for introductory mechanics.

II. THE LABORATORY, CLAIMS, AND EVIDENCE

For several years, the first laboratory that I use in my mechanics class has been based on the standard "half-Atwood's machine" configuration (see [13] for a more complete description). The lab is inspired by a particular sequence of questions on the Force Concept Inventory (FCI) [14]. A low-friction cart is placed on a level track and connected to a string running over a pulley at the end of the track. A mass is affixed at the hanging end of the string. The students are asked to go through the following sequence of tasks:

1. Determine what mass must be hung on the string in order for the cart to move at constant speed when given a gentle push toward the pulley.
2. Keep the mass on the string the same as in task 1, but observe what happens if the cart is given a more substantial push to begin the motion.
3. Double the mass found in task 1 and observe what happens if the cart is given a gentle push.

The carts are equipped with motion detectors that allow the students to automatically generate position vs. time and velocity vs. time graphs.

This lab seemed a useful test bed for the CER rubric. The claims are highly constrained in scope, but are open in the sense that the students are not verifying an accepted value. For task 1, the requested claim is a statement of the mass that was found to work. There is no specific accepted value for this (experimentally it is generally in the range of about 0.5-5 grams for our 550 gram carts). For task 2 the theoretically expected claim from Newtonian mechanics would be that the cart travels at constant speed, but this is not stated in the laboratory. For task 3 the theoretically expected claim would be that the cart will move with steadily increasing speed, but again this is not stated in the instructions. Based on responses to the FCI, it would be reasonable to expect some students to express misconceptions such as the claim that in task 2 the cart should decelerate to the speed seen in task 1 or that in task 3 the cart should travel with a steady speed that is twice as fast as that seen in task 1 [13].

The type of supporting evidence to be used is also constrained and the same type of evidence is useful in all three tasks. In each case either the slope of a position vs. time graph or the value of a velocity vs. time graph could be used to determine whether the speed is steady or changing.

III. REASONING WITH THEORY OR ABOUT EVIDENCE

Before implementing a full version of the CER rubric, for two semesters (Spring 2018 and Fall 2019) I asked students to give their evidence and reasoning without making this an explicit framework for the answer. For example, the prompt for task 3 in the lab was:

> What do you claim should happen if you double the pulling mass that you found in task 1 and again release the cart moving in the direction of the pulley? Explain your reasoning. Collect evidence and present it below. If you were wrong, correct your claim and reasoning.

Assessing student responses demonstrated that reasoning in this laboratory is split into two components. Kuhn *et al.* label these components "theory-based" reasoning and "evidence-based" reasoning [15]. Theory-based reasoning encompasses the application of accepted scientific principles to predict or explain the outcome of a given experimental situation. In terms of the CER rubric, theory-based reasoning would be used to explain why a claim was made and investigated in the first place or how it fits into the larger body of scientific knowledge. Evidence-based reasoning is the ability to interpret data to decide if a particular claim has been supported by the data or not. Evidence-based reasoning is what is requested in the CER's definition. It is a better match to the general education goal of reasoning from evidence. Further, Kuhn *et al.* found that the "most pronounced difference between children and adults is in children's use of theory-based reasoning to justify their inferences" whereas adults "soon begin to attend to the evidence" [15], which emphasizes the importance of helping students to develop evidence-based reasoning strategies. In reality, a good scientist needs both types of reasoning. Scientists must generate and contextualize claims by referencing existing scientific literature in addition to defending those claims with experimental evidence. This duality is reflected in the NGSS. Practice #7 is focused on evidence-based reasoning, but other practices focus more on theory-based reasoning (*e.g.* practice #2, developing and using models) or involve a combination (*e.g.* practice #3, planning and carrying out investigations). The duality is also explicitly mentioned by the authors of the CER rubric. Their full definition of the Reasoning component is [12]:

> "[R]easoning provides a justification for why or how the evidence supports the claim. The reasoning often includes scientific principles or science ideas that students apply to make sense of the data."

An example of a response given by a group of my students that blends theory-based and evidence-based reasoning for task 3 as phrased above is:

> "We believe that the velocity of the cart will increase because the string force is now more than the force of friction. Our claim was correct, the position graph does not have a constant slope, the slope gradually increases. The velocity graph also gradually increases, indicating that the cart was speeding up. This is because the force of the string is now greater than the force of friction."

(The position and velocity graphs provided as evidence are not replicated here, but the students' description is accurate)

The students making this response have demonstrated evidence-based reasoning in the middle two sentences by stating the aspects of the graphed data that support their claim. They have demonstrated theory-based reasoning in the first and final sentences by discussing the force balance for this situation as a justification for their claim.

With the phrasing of tasks 1-3 asking students to "explain your reasoning" as demonstrated in the wording for task 3 above (but not using the full CER framework), I found that I was leading the students toward the use of theory-based reasoning rather than evidence-based reasoning. Evaluating 16 different lab submissions from 2018-19 according to the coding scheme described in Table I yields the data shown in Table II. Uncertainty ranges in Table II resulted from ambiguous responses that described the velocity in the reasoning but did not justify the description by explicit reference to the data (see the "Evidence-based?" code entry in Table I).

In responding to task 1 more than 80% of lab groups employed evidence-based reasoning and only about 20% provided any theory-based reasoning. By task 2, however, less than 70% of the lab groups explicitly demonstrated evidence-based reasoning and over 60% of the responses employed theory-based reasoning in some form. In task 3, 50% or less of the lab groups provided any evidence-based reasoning. I believe there were two reasons for this shift.

TABLE I. Coding scheme used to classify student reasoning. The final category ("Evidence-based?") was multiply coded and gave rise to uncertainty ranges in Table II and Table III.

Reasoning code	Coded statement
Theory-based	mentions force balance
Theory-based	mentions change in a force
Evidence-based	mentions shape of position graph
Evidence-based	references shape or values of velocity graph
Evidence-based?	shows velocity graph and describes velocity without explicit reference to shape (*e.g.* "velocity stays constant", "velocity increased", etc.)

TABLE II. Number and (fraction) of lab reports employing theory-based vs. evidence-based reasoning for tasks 1-3 for the 2018 and 2019 samples (no explicit CER framework). N=16 reports.

Reasoning provided	Task1	Task 2	Task 3
Evidence-based only	12-13 (75-81%)	2 (13%)	2-3 (13-19%)
Both types	1 (6%)	7-9 (43-56%)	3-5 (19-31%)
Theory-based only	2 (13%)	3-5 (19-31%)	6-8 (38-50%)
None provided	0-1 (0-6%)	2 (13%)	2-3 (13-19%)

First is the construction of the tasks. For task 1 there is no way to predict the mass that will balance the friction on the cart and enable it to move with constant speed before data is gathered. The few cases classified as theory-based reasoning in that task were backward-looking justifications amounting to a restatement of the task such as: "Because that force equals out the friction force. That is when the pull force and the friction force in the X direction are balanced." By task 2, however, it is possible to make a legitimate theoretical prediction because the friction force has been measured and balanced in task 1. Students provided theory-based reasoning statements such as: "the friction and force from the string are equal, so the velocity stays constant". Similarly, for task 3 theory-based statements were generally of the form: "the velocity should increase because the tension in the string is greater than the friction".

A second likely reason for the shift to theory-based reasoning is my own unconscious bias. While introducing the lab it was necessary to explain what was meant by "reasoning". Examples were given that explicitly discussed the shape of the graph, and task 1 was the natural focus of this introduction. In later tasks no explicit examples of reasoning were given, but in answering student questions about the tasks I tended to discuss the Newtonian principles that apply, which could have led students to incorporate those concepts into their reasoning. The structure of the tasks as written in 2018-19 confirms a predisposition on my part to prioritize theory-based reasoning. Task 3 as written above transitioned directly from claim to reasoning: "What do you claim should happen if you double the pulling mass that you found in task 1 and again release the cart moving in the direction of the pulley? Explain your reasoning." Only then were students asked to collect evidence. This put the CER concept out of order. In effect the prompt implied that students should be able to reason through the task without the evidence, leading them to employ theory-based reasoning. That expectation was contrary to the purpose of a discovery lab and, as seen in the responses of my students, undermined the teaching and assessing of evidence-based reasoning. This led to a change of approach.

IV. THE CER FRAMEWORK

In 2020 and 2021 I administered a revised version of the laboratory that explicitly used the CER framework. In task 1 I also gave the students some explicit guidance about what the Evidence and Reasoning components should involve. In the revised lab, after task 1 was introduced, students were given the following explicit prompts:

> Task 1 Claim (what mass is needed):
> Task 1 Evidence (graphs supporting claim):
> Task 1 Reasoning (explain in words how your evidence supports the claim):

Later tasks also included an explicit CER prompt, and the wording of the introductions to the tasks was revised to better reflect the structure of the CER. In the 2020-21 version of the lab task 3 was written:

> What do you claim should happen if you double the pulling mass that you found in task 1 and again release the cart moving in the direction of the pulley? Make a prediction and enter it below. Then, collect evidence, make a final claim, and explain your reasoning.
> Task 3 Initial Claim:
> Task 3 Final Claim:
> Task 3 Evidence:
> Task 3 Reasoning:

This new wording indicates that evidence should be collected before the reasoning is established. Evaluating 12 lab reports from 2020-21 using the coding scheme of Table I yields the data shown in Table III. With the explicit CER framework and revised wording for the tasks, the significant majority of lab groups provided evidence-based reasoning for all three tasks.

V. FUTURE GOALS: ASSESSING AND TEACHING EVIDENCE-BASED REASONING

Having found that the explicit CER framework can help both the students and me as the instructor focus on evidence-based reasoning in the laboratory, the next step is to consider

TABLE III. Number and (fraction) of lab reports employing theory-based vs. evidence-based reasoning for tasks 1-3 for the 2020 and 2021 samples (explicit CER framework). N=12 reports.

Reasoning provided	Task1	Task 2	Task 3
Evidence-based only	12 (100%)	10-11 (83-92%)	6 (50%)
Both types	0	0-1 (0-8%)	3-6 (25-50%)
Theory-based only	0	0-1 (0-8%)	0-3 (0-25%)
None provided	0	0-1 (0-8%)	0

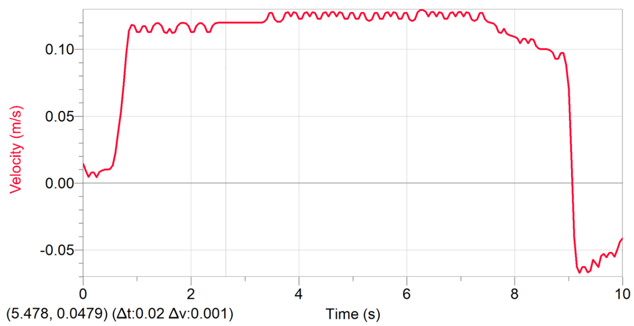

FIG. 1. A typical velocity vs. time graph from task 1.

how best to assess and teach this skill. In assessing student reasoning, one problem is deciding the correct balance between requiring explicit details so that the reasoning is clear versus assuming a standard knowledge base so that the explanations are not overly tedious to create or to read.

As an illustrative example of this issue, consider the velocity vs. time graph shown as Fig. 1, which is typical of the evidence presented for tasks 1 and 2 of the lab. The student who provided that graph stated the reasoning: "Velocity of the cart remains at a constant level...". To someone familiar with interpreting graphical data but unfamiliar with this particular lab that statement would seem absurd. It only makes sense provided that I as the instructor assume that the student knows the sharp rise occurring at a time of about 1s is due to the initial push of the cart and the sharp drop at around 9s is the result of the collision with the pulley at the end of the track. This lab requires students to focus on the region where the cart is moving under the influence of the string tension and friction only (in the horizontal direction).

Only a few students provided quantitative statements about the time range considered or the velocity obtained. In the 2020-21 sample four reports out of 12 (33%) gave a specific time range considered for any of the tasks. Three reports out of 12 (25%) gave a specific velocity value for any of the tasks. Requiring this type of quantitative reasoning could help the instructor reduce the number of assumptions made when assessing the reasoning statements. Such a requirement would have reduced the ambiguity in coding responses for this work.

This raises the question of whether requesting quantitative reasoning could also be beneficial as a teaching tool. The lab was largely successful at prompting students to arrive at Newtonian claims for the three tasks. For task 2, of 30 total lab reports evaluated from 2018-2021, nine reports (30%) stated a non-Newtonian initial claim and seven of the nine corrected the claim after evidence was gathered. For task 3 three reports out of 30 (10%) initially stated a non-Newtonian claim, and one of those three corrected the claim after evidence was gathered. For the four instances of a report providing reasoning to support a non-Newtonian claim, it appears that requesting a quantitative discussion of the velocity in the reasoning would have been helpful.

As an example, consider one of the lab groups that attempted to support a non-Newtonian claim for task 2. In task 1, the group provided this reasoning statement:

> "The velocity graph remains constant at 0.392 m/s from time 1.3 seconds-2.3 seconds..."

which is an accurate description of their graphed data and displays a sufficient level of quantitative reasoning without being tedious. For task 2, their claim was:

> "Our initial prediction was that our cart would slow down due to the fact that the cart can't accelerate faster than the mass of the paper clips (0.8 grams in total) was pulling on it."

A quantitative description of their task 2 graph following the format of their own task 1 reasoning would have been to say that their velocity graph remained constant at approximately 0.82 m/s from time 0.8 seconds-1.2 seconds. That reasoning would have invalidated their initial prediction. The reasoning they actually provided was:

> "We believe our prediction was right because there is a slight decrease in the velocity, but it is almost an intangible factor with the amount of track that we're given."

It is worth investigating whether requiring students to provide a quantitative description of their data could help prevent this type of confirmation bias in their reasoning.

VI. CONCLUSION

If physics labs are to fulfill the general education goal of teaching evidence-based reasoning then this must be an explicit component of the lab report. The CER framework is one possible approach to accomplish this. I have adapted one lab for introductory mechanics to use the CER framework. In doing so, I found that the CER framework does help to encourage both the students and the instructor to explicitly report evidence-based reasoning as opposed to falling back on theory-based reasoning alone. The framework itself is not enough, however. It must be supported by appropriate framing of the lab tasks. In order to teach the reasoning step in early labs it is probably also helpful to specify that students must explicitly reference numerical data in their reasoning statement.

[1] https://www.nextgenscience.org. Retrieved 4/14/2022.
[2] National Research Council, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Academies Press, Washington, 2012).
[3] N.G. Holmes and D.A. Bonn, "Quantitative comparisons to promote inquiry in the introductory physics lab", Phys. Teach. **53**, 352 (2015).
[4] P.W. Laws, "Calculus-based physics without lectures", Phys. Today **44**(12), 24 (1991).
[5] L.C. McDermott and the Physics Education Group at the University of Washington, *Physics by Inquiry, Vol. 1* (Wiley, 1995).
[6] M. Wells, D. Hestenes and G. Swackhamer, "A modeling method for high school physics instruction", Am. J. Phys. **63**, 606 (1995).
[7] E. Etkina and A. Van Heuvelen, "Investigative Science Learning Environment – A science process approach to learning physics", in *Research-Based Reform of University Physics*, edited by E.F. Redish and P.J. Cooney (American Association of Physics Teachers, College Park, 2007), Reviews in PER Vol. 1, http://www.per-central.org/document/ServeFile.cfm?ID=4988. Retrieved 7/15/2022.
[8] R.J. Beichner, J.M. Saul, D.S. Abbott, J.J. Morse, D.L. Deardorff, R.J. Allain, S.W. Bonham, M.H. Dancy and J.S. Risley, "The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project", in *Research-Based Reform of University Physics*, edited by E.F. Redish and P.J. Cooney (American Association of Physics Teachers, College Park, 2007), Reviews in PER Vol. 1, http://www.per-central.org/document/ServeFile.cfm?ID=4517. Retrieved 7/15/2022.
[9] E.F. Redish and D. Hammer, "Reinventing college physics for biologists: Explicating an epistemological curriculum", Am. J. Phys. **77**, 629 (2009).
[10] N.G. Holmes, *Structured Quantitative Inquiry Labs: Developing critical thinking in the introductory physics laboratory*, Ph.D. thesis, University of British Columbia, 2014.
[11] K.L. McNeill and J. Krajcik, *Supporting Grade 5-8 Students in Constructing Explanations in Science: the Claim, Evidence and Reasoning Framework for Talk and Writing* (Pearson Education, Boston, 2011).
[12] K.L. McNeill and D.M. Martin, "Claims, evidence, and reasoning", Science and Children **48**, 52 (2011).
[13] A. Pawl, "Using Force Concept Inventory data to develop impactful class activities", Phys. Teach. **58**, 94 (2020).
[14] D. Hestenes, M. Wells and G. Swackhamer, "Force concept inventory", Phys. Teach. **30**, 141 (1992).
[15] D. Kuhn, M. Garcia-Mila, A. Zohar and C. Andersen, "Strategies of Knowledge Acquisition", *Monographs of the Society for Research in Child Development* **60**(4), (1995).