A comparison of students' written explanations and CLASS responses

Gabriel Rowe¹, and Anna McLean Phillips²

¹*Clark University, 950 Main Street, Worcester, MA, 01610*

²*Department of Physics and Astronomy, Tufts University, 574 Boston Avenue, Medford, MA, 02155*

We examined students' responses on the Colorado Learning Attitudes about Science Survey (CLASS) and their homework within a freshman physics course. The homeworks were designed to elicit detailed responses and sense-making. Students completed the CLASS prior to the start of the course. Student responses on the first two homeworks of the semester were coded for instances of evidence of sense-making, such as checking if an answer is reasonable and providing multiple explanations. We had matched CLASS responses and homeworks for 26 students. Our results show that evidence of sense-making on the homework assignments may be uncorrelated with percent favorable response on the CLASS and weakly correlated with percent unfavorable response. These results suggest that students' formal beliefs about physics, as measured by the CLASS, may be only weakly related to how students engage in solving physics problems. Given the small sample size of our work, further study is needed.

I. INTRODUCTION

Within the Physics Education Research (PER) community, surveys such as the Colorado Learning Attitudes about Science Survey (CLASS) [1] have been used to measure students' beliefs about physics for some time. The CLASS is used to measure how well students' professed beliefs about physics line up with the beliefs of physicists, which are frequently referred to as "expert-like." These surveys have been used to assess how courses at dozens of institutions have shaped the beliefs of thousands of students, a matter of great interest to the PER community [2]. In this study, we sought to compare how these beliefs, as measured by CLASS, correspond to what students actually do while solve physics problems.

Researchers in science education have argued for distinguishing between the beliefs about science that students describe in surveys and interviews, which are referred to as formal or declarative beliefs, and the beliefs that underlie what they actually do when they engage in doing science, which are referred to as practical beliefs. Sandoval describes how individuals' ideas about how they can construct knowledge through science inquiry (their "practical epistemology") may not be closely related to their ideas about how knowledge is structured and constructed in professional science (their "formal epistemology") [3]. This distinction, he argues, explains why the decisions students make while engaging in inquiry activities may be only weakly related to how they report their understanding of science when directly asked on surveys and in interviews. Similarly, Salter and Aktins' recent work in a science inquiry course demonstrates how students' declarative statements about the Nature of Science (NOS) on surveys do not reliably relate to students' practices when they engage in science inquiry [4].

In light of this work, we asked how well students' formal beliefs, as reported on the CLASS, align with their written explanations of homework problems. One might expect that a high percentage favorable response on the CLASS would cor-

respond to certain behaviors when completing physics problems. For example, a student who agrees with the statement "When I am solving a physics problem, I try to decide what would be a reasonable value for the answer" could be more likely to include a statement about whether or not their answer makes sense in their homework response.

To begin to answer these questions, we used data from an introductory physics class at Tufts University. Students were asked to complete the CLASS prior to the first day of class. The homework assignments in the course were designed to elicit detailed responses and encourage sense-making. We coded student responses for various features that appeared related to whether or not a student was engaging in "expert-like" thinking or sense-making about the problem.

We found little relationship between students' responses on the CLASS and these features, both at the level of individual codes and questions and at the level of overall scores. We have 26 matched sets of homework and CLASS responses; therefore, our results are only preliminary. However, these findings suggest that the distinctions made in science education research between formal and practical beliefs should inform our interpretations of responses on the CLASS. It is possible students' formal beliefs as measured by the CLASS scores may not be closely related to students' engagement in doing physics. Therefore, the PER community may need to give further attention to the interpretation of CLASS as a measure of how well courses have met instructional goals.

II. METHODS

The course that we drew data from is similar in approach to that described by Redish and Hammer, in that class discussions, homework assignments, and laboratories were specifically designed to shape students' beliefs about physics [5]. As with several sections of the course described in Ref. [5], Hammer was the instructor, though this course was significantly smaller and took place at a different university. Weekly

homework assignments consist of 3-5 problems, and they are graded primarily on whether or not the response is sensible rather than on correctness. The instructions on the first homework assignment include language emphasizing this:

Here's the really important thing: The first priority is that the TA can understand your reasoning—what you write needs to *make sense*. In this, I ask them to imagine they don't understand how to answer the question, and you're going to explain your reasoning to them. **Can they follow what you're saying?** If they can't, they can't give you credit for your solution, whether or not your answer is correct. If they can, if they can follow your reasoning and it makes sense, that's worth points. (emphasis original)

There were approximately 50 students enrolled in the course, most of whom were freshman in the school of engineering. Due to enrollment fluctuations during the first two weeks of the course, as well as a relatively low (approx. 60%) response rate on the CLASS, we have matched homework assignments and valid CLASS responses for 26 students. Thus, there is a selection bias towards students who signed up for the class and did not switch courses during add/drop period.

A link to complete the CLASS was emailed to students registered for the course a week prior to the start of the course, and students were asked to complete the CLASS prior to the first class meeting. Students were offered a small amount of participation credit for going to the survey and completing the study consent form. Responses after the first class period were not counted.

Student responses on the first two homework assignments were analyzed in order to capture how students solved problems near the beginning of the course. Though the second homework assignment was due a week and a half into the class, we chose to include it to have a larger number of problems to analyze for each student. We only coded the homework assignments of students for whom we had valid CLASS responses and both homework assignments.

We developed codes that captured evidence of expert-like thinking and sense-making. The goal was to generate sub-scores that could be compared to individual CLASS questions, clusters of related CLASS questions, and an overall score that could be compared to the overall favorable and unfavorable scores on CLASS.

To generate codes the codes, we first read through all of the homeworks and noted common features of “good” answers—answers where we had the impression that students were providing clear explanations that were connected to the physics the problem was modeling. These features were condensed into codes 1-8 in Table I. For example, one student explained how they arrived at their answer through a collaborative process with friends where they ran a few tests on their theories. While their answer was not correct, this is evidence that they were sense-making around the question. This was coded for

code 5 (mentions an at home experiment) and code 4 (references physical interactions between objects). One might expect such a statement to correspond to favorable responses to CLASS questions such as “The subject of physics has little relation to what I experience in the real world” and “To understand physics I discuss it with friends and other students.”

We also noted features of answers that provided no evidence that the students were sense-making, which were used to create codes 9-11. For instance, one student responded that a ball increasing speed discretely at 10cm/s (see Fig. 1), would have traveled over 14 meters after only 6 seconds. This coded for code 11: Gives an unreasonable answer. One might expect that such responses would be related to unfavorable responses to CLASS questions such as “When I am solving a physics problem, I try to decide what would be a reasonable value for the answer.”

We generated tentative descriptions of each code, and a subset of student homeworks ($n=4$) were coded by each author. We discussed disagreements and refined the codes, including noting that alternative answers and explanations should only be coded if a student went above and beyond what was explicitly requested in the problem statement. We also noted the difficulty in coding one particular problem of homework set 1: the majority of the codes did not apply and there was some variation in the number of physical interactions between objects (code 5; see Table I and Fig. 2) each author identified. Because of this, we chose not to include this problem in our analysis. Therefore, our analysis includes a total of 7 problems for each student. A list of the codes is in Table I. A sample homework problem is in Fig. 1, with sample coded student responses in Figs. 2 and 3.

Both authors independently coded the two problems on the first homework assignment. The inter-rater reliability rate was 93%. We discussed disagreements and arrived at consensus for the final codes. Given the high inter-rater reliability rate, only Rowe coded the second homework assignment.

In order to create a quantitative measure that we could use in our analysis, we assigned scores of +1 or -1 to each code. For each student, we summed the scores for each problem within each code to create sub-scores for each code. We also summed across the codes to create an overall score for each student. We compared student sub-scores for particular codes to the related CLASS questions listed in Table I both individually and as a cluster of questions. We calculated Pearson correlation coefficients and p-values for each. Code 8, “Uses a picture or diagram as a tool” and code 9 “Includes no explanatory text” did not neatly map onto particular CLASS questions. We identified code 8 as common within responses where students seemed to be making sense of the physical scenario. Code 9, which was applied rarely, was created to mark responses that provided us with little to no evidence of students’ reasoning about a question.

We also compared the overall homework score for each student to their percent favorable and unfavorable responses on the CLASS. We calculated Pearson correlation coefficients and p-values for each.

1. I hit that bowling ball from lecture with the mallet, starting from rest. Suppose every time I hit it, its velocity increases by exactly 10cm/s . You start timing, $t=0$ seconds, at the instant I hit the ball the first time, and I hit it once every second. Two simplifying approximations: (1) Suppose each hit is very quick, so the speed changes “instantly.” (2) Suppose the ball doesn’t slow down between hits.
- At $t = 6$ seconds, how far has the ball moved from its starting position?
 - Write an expression for how far the ball has moved after t seconds.
 - Now I change my technique. I ease up how hard I hit the ball, so that now it speeds up 5 cm/s with every hit. But now I hit it once every half second. Start the clock again at the instant I hit the ball the first time, starting from rest, and again say how far the ball has moved at 6 seconds.

FIG. 1. Problem 1 from the first homework.

TABLE I. Codes used to analyze homework.

Code	Scoring	Related CLASS Questions
1. Explains an expression	+1 for each occurrence	8, 13, 24, 32, 39
2. Checks if the answer makes sense	+1 for each answer checked	2, 13, 23, 35, 37, 39
3. Checks math	+1 for each answer checked	13, 23, 36
4. References physical interactions	+1 for each interaction	35, 37, 42
5. Mentions an at-home experiment	+1 for each experiment mentioned	35, 37
6. Gives alternative answers	+1 for each alternate answer beyond the requirement	15, 36
7. Gives multiple explanations for a particular answer	+1 for each explanation beyond the requirement	10, 15, 21, 36
8. Uses a picture or diagram as a tool	+1 for each type of diagram	N/A
9. Includes no explanatory words	-1 if no explanatory words	N/A
10. Uses math that does not model the physical scenario	-1 for each non-physical model used	8, 13, 23, 24, 35
11. Gives an unreasonable answer	-1 for each unreasonable answer	2, 8, 23, 35

FIG. 2. Student A’s response. For this response, we coded for “references physical interactions,” “explains an expression,” “uses a diagram or picture,” and “checks math” once each. Note that the same physical interaction is mentioned twice and is only coded once. The total score for this problem is +4.

III. FINDINGS

We found essentially no correlation between students’ percent favorable response on the CLASS and their overall homework score. The Pearson correlation coefficient is $r=0.17$ and the associated p -value is 0.40. There was a weak relationship between students’ percent unfavorable response on the CLASS and their overall homework score. The Pearson correlation coefficient is $r=-0.33$, with p -value 0.10. Scatterplots are shown in Fig. 4. Comparisons between sub-scores and clusters of related CLASS questions did not yield any r -values larger than 0.2. For sub-scores on three codes,

one question each was correlated with r -value ≥ 0.3 :

- Code 1 and question 8, “When I solve a physics problem, I locate an equation that uses the variables given in the problem and plug in the values” ($r=0.42$, $p=0.03$);
- Code 7 and question 15, “If I get stuck on a physics problem on my first try, I usually try to figure out a different way that works” ($r=0.31$, $p=0.12$); and
- Code 10 and question 23, “In doing a physics problem, if my calculation gives a result very different from what I’d expect, I’d trust the calculation rather than going back through the problem” ($r=0.50$, $p=0.01$).

FIG. 3. Student B's response. In this response, we coded for "uses a picture or diagram as a tool" once. Though there are two graphs, they are the same type of graph (velocity versus time) and they are used in the same way. This student includes "no explanatory text," resulting in a score of -1 for that category. They also use mathematics that does not reflect the physical scenario, which results in a -1 score in "non-physical math." The total score for this problem is -1.

FIG. 4. Students' overall homework score vs % favorable and %unfavorable responses on CLASS

IV. DISCUSSION AND CONCLUSION

Our results are limited by a relatively small sample size, both in numbers of students and in the number of problems analyzed for each student. Furthermore, while analyzing student work provides us with an artifact of their thinking, students may engage in additional sense-making thought processes that they do not report while writing their problem set. This could easily apply to cases where students provided no explanatory text whatsoever. While directly prompting to provide thorough explanations and grading on the sensibility of those explanations may encourage detailed responses, we admit that the written responses do not provide us with a full picture of how a student solved the problems. These homework assignments encouraged specific types of explanations and behaviors; it is probable that very different results would

be seen in courses where significant written explanations are not required. It is possible that students may engage in sense-making when they are presented with problems such as these that encourage them to think about the physical world, even if their responses in CLASS indicated a novice-like view.

Despite these limitations, our work raises questions about how well CLASS and similar surveys reflect how students think about physics when they approach physics problems. While we see some alignment on several particular questions and codes, we do not see significant relationships between overall evidence of expert-like thinking and sense-making on homework responses and CLASS favorable and unfavorable responses. Given this lack of alignment between CLASS responses and homework responses, shifts in students' formal beliefs, as seen studies using CLASS, may only be loosely related to shifts in students' actual practices when solving physics problems. Our work also suggests that work in science education on the difference between formal and practical beliefs should inform the use and interpretation of surveys such as CLASS. As our study only examined student responses at the start of the semester and included a very small N value, further work is needed to examine the relationship between shifts over the course of the semester.

ACKNOWLEDGMENTS

We would like to thank David Hammer for facilitating data collection in his course, Jessica Watkins for her assistance with data collection and her feedback on a draft of this paper, Sherry Cook at The Clark School for arranging Gabriel's internship, and Jennifer Radoff for her feedback on drafts of this paper.

- [1] W. Adams, K. Perkins, N. Podolefsky, M. Dubson, N. Finkelstein, and C. Wieman, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006).
- [2] A. Madsen, S. B. McKagan, and E. C. Sayre, Phys. Rev. ST

- Phys. Educ. Res. **11**, 010115 (2015).
- [3] W. A. Sandoval, Sci. Educ., **89**, 4 (2005).
- [4] I. Y. Salter and L. J. Atkins, Sci. Educ., **98**, 1 (2014).
- [5] E. F. Redish, and D. Hammer, Am. J. Phys., **77**, 7 (2009).