Are Students' Responses and Behaviors Consistent?

Umporn Wutchana[1], Narumon Emarat[2] and Eugenia Etkina[3]

[1]*Institute for Innovative Learning, Mahidol University, Bangkok, 10400, Thailand.*
[2]*Department of Physics, Faculty of Science, Mahidol University, Bangkok, Thailand*
[3]*Graduate School of Education, Rutgers University, New Brunswick, New Jersey 08901, USA*

Abstract. In our project we attempted to determine if students whose beliefs about physics are more expert-like and less expert-like, as judged by the CLASS survey, are different in terms of their approaches to learning physics and whether their behaviors in the classroom are consistent with their responses to the surveys. All students enrolled in the second semester of an introductory physics course took the CLASS survey. We used survey results to identify expert-like and non-expert like students to participate in the study. We selected four highest scoring and four lowest scoring students. We then observed those students in laboratories and problem-solving recitations during one semester and interviewed them at the end. We found some inconsistencies between students' responses to the survey and their actual behaviors as well as several significant differences in behaviors of more expert-like and less expert-like students. This work was supported by The institute for the Promotion of Teaching Science and Technology, Thailand.

Keywords: Beliefs, Self-efficacy, CLASS, Survey
PACS: 01.40.Fk, 01.40Ha

INTRODUCTION

Multiple studies investigate students' beliefs about physics, such as the structure of physics knowledge, connections between physics and the real world; and their beliefs about learning physics. [1-3] One can assess approaches to learning through observations and beliefs through interviews. In physics education research beliefs are often assessed through surveys, such as Maryland Physics Expectations Survey (MPEX) [4], the Epistemological Beliefs Assessment for Physical Science (EBAPS) [5], and the Colorado Learning Attitudes about Science Survey (CLASS). [6] However, it is not clear whether the assessment of students' beliefs based on the surveys matches their actual approaches to learning.

When students take a survey, they are asked to rate their agreement with the responses such as: "I study physics to learn knowledge that will be useful in my life outside of school," or "If I get stuck on a physics problem on my first try, I usually try to figure out a different way that works." These ratings occur outside of normal course activity and are self-reported.

Is it possible that students' responses do not match their actual approaches to learning? Some students might think that they function in one fashion, but actually behave differently. Students may answer survey questions in a favorable fashion, but have unfavorable behaviors in the classroom. At the same time it is possible that students who do not believe that they are good in physics will choose unfavorable responses, but in the classroom they behave differently. As one study shows, most students who display novice beliefs about physics and learning of physics are, in fact, quite aware of what physicists think on this matter; they just do not believe that these ideas are valid, relevant, or useful for them [7]. Thus, if we wish to use these surveys as a measure of classroom learning, it is important to study the relationship between student responses to the surveys and their actual behaviors – the subject of the present paper. Specifically, we intend to answer the following questions: (1) Do students with extreme low and high scores on the survey behave differently from each other? (2) Are students' classroom behaviors consistent with their scores on the survey?

DATA COLLECTION

Our project was conducted in the second semester of an introductory physics course for science majors with approximately 180 students. There were two 55-min lectures per week, one 80-min problem solving session, and a 3-hour lab. The course followed the Investigative Science Learning Environment curriculum [8]. In "lectures" students worked actively to help develop and apply concepts, in recitations they solved multiple representation problems, and in labs they designed their own experiments being guided by scientific abilities rubrics [9].

In this study, we used four primary data sources to allow for triangulation of data. Our first data source was students' responses to the CLASS survey, which all students enrolled in the course took at the beginning of the second semester. The survey consists of 42 statements to which students respond using a 5-point Likert scale. The 'Overall' favorable score is measured as the average percentage of statements to which a student answers similarly to an expert physicist. The second and third sources of data consisted of field notes collected during observations of selected students in labs and in recitations. The students were selected according to the following procedure. After all students took CLASS, we made a list of all the students who had extremely high and low scores on the survey. We then selected eight of them, four highest scoring and four lowest scoring students, from different lab and recitation sections. The students' scores on the CLASS and their exam scores are presented in Table 1 (there is no correlation between their CLASS responses and exam scores). We then observed those students in laboratories and recitations during the whole semester. The observations focused on one student in the group and thus allowed us to observe her/his behavior in detail. We collected detailed field notes of everything that each student in the study did and said during the classes. After week 1, we analyzed the notes and developed a provisional coding scheme using existing approaches to coding student behaviors [10, 11]. The next week, the codes were modified to match new emerged student behaviors. After the second week all observed behaviors fit into the developed scheme. Two coders achieved an acceptable reliability of the scheme by coding about 15% of the notes. The coding descriptions are presented in Table 2.

As a fourth data source, we interviewed the eight students at the end of the semester. The interview protocol was constructed based on the CLASS questions and the observation-coding scheme. The goal of the interviews was to determine whether there was a consistency between what students reported and what they actually did in the classroom. The students were asked about their experiences in lectures, labs and recitations. All student interviews were audio taped and then transcribed. The same coding scheme was used to code the interview transcripts.

TABLE 1. Favorable responses on CLASS and Exam scores

	Low Scorings				High Scorings			
Students' Scores	1	2	3	4	5	6	7	8
Fav (%) on CLASS	17	28	28	31	75	75	81	86
Exam Scores (%)	83	72	90	70	95	68	77	90

TABLE 2. Coding Scheme

Codes	Description
Coherence	Discussing related physics concepts or physical quantities; relating to previous experience i.e. previous lab; lecture etc.; connecting to real life experience.
Metacognition	Global planning; monitoring.
Persistence	Designing another independent experiment for checking the experimental results or testing a hypothesis; repeating the experiment, trying to find another solution.
Reasoning	Constructing an explanation; providing a mechanistic explanation; demonstrating reasoning with multiple representations (including mathematics).
Leadership/Confidence	Assessing others; teaching others; guiding others (showing leadership); multitasking.
Authority	Asking students in other groups for clear directions; asking a TA's approval or for clear directions
Scientific Abilities	Discussing experimental design; setting a hypothesis for testing; making a prediction of the outcome of the experiment; discussing assumptions, uncertainties, experimental results, etc.

FINDINGS

The first research question relates to student behaviors. First we noted that on average low-scoring CLASS students spoke less in class than higher scoring: for labs, low scoring students have 125 ± 51 coded statements and high scoring students have 212 ± 52 coded statements (significant at α=0.02), for recitations respective numbers are 46 ± 23 and 77 ± 14 (close to being significant, α=0.09). Figure 1 shows descriptive results of coded student behaviors. The heights of the bars represent the median percentages of coded statements that belong to a particular coded category. Students with high scores on CLASS show higher percentages of Scientific Abilities statements and slightly higher percentages of Coherence and Persistence statements while the students with lower scores show a higher percentage of Metacognition statements. However, these differences are not statistically significant. Both groups show a higher percent of Reasoning statements in recitations than in labs. The major differences between high and low scoring students in terms of their behaviors are in Leadership/Confidence and Authority. The high CLASS scoring students have a higher statement percent coded as Leadership/Confidence and fewer as Authority. Table 3 shows the results of Mann-Whitney U test used to determine Statistical significance. In the labs, higher scoring students have significantly higher percentages for Leadership/Confidence and lower scoring students have significantly higher percentages for Authority. These differences appear again in recitations. The significant difference is in Leadership/Confidence. Although in Authority the difference is not statistically significant, it is close to being significant. High and low scoring students have no significant difference in terms of their Scientific Abilities, Coherence, Metacognition, Persistence and Reasoning behaviors. The major differences are only in Leadership/Confidence and Authority.

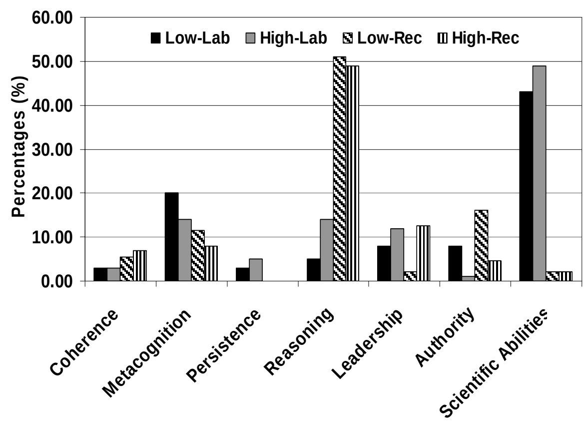

FIGURE 1. Median percentages (%) of coded statements of high and low scoring students in labs and recitations.

TABLE 3. Comparison of high and low CLASS scoring students' classroom behaviors.

		Labs		Recitations	
	Groups	Mean Rank	Sig.*	Mean Rank	Sig.*
Coherence	Low	7.8	0.16	8.0	0.67
	High	11.2		9.0	
Metacognition	Low	10.6	0.37	9.3	0.49
	High	8.4		7.7	
Persistence	Low	8.1	0.23	8.0	0.32
	High	10.9		9.0	
Reasoning	Low	7.8	0.18	9.3	0.53
	High	11.2		7.8	
Leadership / Confidence	Low	6.9	0.04*	5.3	0.006 **
	High	12.1		11.8	
Authority	Low	12.9	0.006 **	10.6	0.07
	High	6.1		6.4	
Scientific-Abilities	Low	8.94	0.66	7.69	0.49
	High	10.06		9.31	

* Significant 2-tailed

TABLE 4. Average percentages of coded statements of high and low CLASS scoring students' classroom behaviors, self-assessments, and their CLASS scores in each category.

	Groups	Labs	Recs	Inter-view	CLASS
Coherence	Low	3.1	6.5	30	19
	High	5.6	8.4	34	81
Persistence	Low	3.6	0.0	29	40
	High	7.0	0.5	38	81
Reasoning	Low	8.3	50.6	16	41
	High	13.1	48.3	17	85
Leadership / Confidence	Low	7.6	2.5	-	24
	High	12.4	11.3	-	81
Authority	Low	7.8	14.8	26	47
	High	2.1	5.0	12	38

These results indicate that the students with high and low CLASS scores actually behave rather similarly except for the amount of talk and Leadership/Confidence and Authority behaviors.

Are students' behaviors consistent with interview responses and CLASS responses? To answer this question we compared the average percentages of coded statements for labs, recitations and interviews for different groups (Table 4). We picked CLASS survey items to match our coding scheme. The table presents CLASS favorable responses (% of max score) for all categories except Authority for which we show students' unfavorable responses instead.

In Coherence cluster, high scoring students' CLASS scores are four times higher than those of low scoring students. Thus, we expect to see much more Coherence in their behaviors than in the behaviors of low scoring students. The results show that high scoring students do have slightly higher percentage of Coherence-coded statements but not drastically different from low scoring students. Even in the interviews, the two groups of students are much closer to each other than on CLASS. We found similar results for Persistence. High scoring students show nearly two times more Persistence than low scoring students in labs (this feature does not appear in recitations), but their interview responses are not very different. In terms of Reasoning the high scoring students do better in labs but slightly worse in recitations. Interviews show no difference.

We do not have data from interviews for Leadership/Confidence. However, the results from labs and recitations show that high scoring students have more Leadership/Confidence. The percentages are nearly two times in labs and four times in recitations, and are consistent with the survey results.

In contradiction, both groups of the students do not have much difference on CLASS scores in terms of their Authority. However, the results show that the lower scoring group has much higher percentages of

Authority statements than the higher scoring group in labs, recitations, and interviews.

Based on our data we can say that students' behaviors are inconsistent with CLASS when they self-reported Coherence, Persistence, Reasoning and Authority. However, we found consistency between students' self-reported and actual classroom behaviors in Leadership/Confidence. High scoring students show higher Leadership/Confidence on all measures.

DISCUSSION

We found that although high and low scoring CLASS students chosen for the study had similar physics grades, lower scoring students talked significantly less in class. We also found that high and low scoring students had no significant difference in term of their Scientific Abilities, Coherence, Metacognition, Persistence and Reasoning behaviors. The major differences were only in Leadership/Confidence and Authority. The high scoring students have higher percentages for Leadership/Confidence and smaller Authority percentages than the lower scoring students. Students with high Leadership/Confidence were assessing others, teaching others, guiding others (showing leadership) and multitasking during the classes. Could the observed differences in the amount of talk and CLASS responses be explained by the differences in student self-confidence or their belief about how capable they are? It is reasonable that more confident students are more inclined to speak in class and might ask TAs fewer questions or need less confirmation. On the other hand, the less confident students might try to ask both TAs and friends to check their understanding.

A concept of self-efficacy might be helpful here [12]. People with high self-efficacy, those who believe they can perform well, are more likely to view difficult tasks as something to be mastered rather than something to be avoided. Students with different self-efficacy are likely to respond to the survey differently. Here are examples of CLASS items: If I get stuck on a physics problem on my first try, I usually try to figure out a different way that works; I do not spend more than five minutes stuck on a physics problem before giving up or seeking help from someone else; If I get stuck on a physics problem, there is no chance I'll figure it out on my own. A student with high self-efficacy might agree with the first statement and disagree with the two latter statements. A high-efficacy person believes that she/he can figure it out on her/his own with just trying in a different way and spending more time. A lower self-efficacy student might disagree with the first statement and agree with the latter statements. Low-efficacy person does not believe that she/he can figure it out without help.

That means self-efficacy might affect how one responds to the survey and cause inconsistencies between students' survey responses and their behaviors. High scoring students might feel more comfortable with the survey because of their self-efficacy. Low-efficacy students do similar things in class but have low confidence to say that they do.

SUMMARY

Focusing on students with extreme high and low CLASS scores, we found the major differences between extremely high and low scoring students' behaviors. The high scoring students speak more and demonstrate more Leadership/Confidence and less Authority behaviors than the lower scoring students. Moreover, the observations indicate that students behave differently from what they report on the survey. Thus it might be difficult to access students' beliefs about physics and learning physics by using surveys only.

There are several limitations to our study. First we had a small number of students; secondly, we did not videotape them and used only field notes for the analysis of their behaviors. Both factors limit the generalizability and the reliability of our findings.

REFERENCES

1. K. K. Perkins, W. K. Adams, S. J. Pollock, N. D. Finkelstein, and C. E. Wienman, *PERC Proceeding*, 61-64, (2004).
2. L. Lising, and A. Elby, *Am. J. Phys.*, **73**(4), 372-382, (2005).
3. G. Kortemeyer, *Phys. Rev. ST Phys. Edu. Res.*, **3**(1), 010101, (2007).
4. E. F. Redish, J. M. Saul, and R. N. Steinberg, *Am. J. Phys.*, **66**(3), 212-224, (1998).
5. A. Elby, http://www2.physics.umd.edu/~elby/EBAPS/home.ht
6. W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, *Phys. Rev. ST Phys. Educ. Res.*, **2**(1), 010101, (2007).
7. K. E. Gray, W. K. Adams, C. E. Wieman, and K. K. Perkins, *Phys. Rev. ST Phys. Educ. Res.,* **4**(2), 020106, (2008).
8. E. Etkina and A. Van Heuvelen, *in Research Based Reform of University Physics*, (2007).
9. E. Etkina, A. V. Heuvelen, S. White-Brahmia, D. T. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and A. Warren, *Phys. Rev. ST Phys. Educ. Res*. **2**(2), 020103, (2006).
10. D. B. May, and E. Etkina, *Am. J. Phys.*, **70**(12), 1249-1258, (2002).
11. A. Karelina, , and E. Etkina, *Phys. Rev. ST Phys. Educ. Res.* **3**(2), 020106, (2007).
12. A. Bandura, *Psychol. Rev.*, 84(2), 191-215, (1987).