

Student Background, Attitudes, and Performance in Physics Breadth Courses

Author: Rikki Landau

Date: August 23rd, 2013

Abstract

This report focuses on how physics breadth courses at the University of Toronto change student attitudes about physics. Students in *PHY100H1S* and *PHY205H1S* were given the Colorado Learning Attitudes About Science Survey (CLASS). Correlations between CLASS scores, course grades, and students' programs of study were investigated. These results may inform course instructors of issues that non-physics majors have in understanding physical concepts.

<u>Table of Contents</u>	<u>Page Number</u>
1. Introduction	2
2. CLASS Literature Review	3
2.1 The development of CLASS v. III	3
2.2 Uses and Results of the Administration of CLASS	5
3. Methods	9
4. Results and Discussion	12
4.1 Final Course Grades	12
4.2 Gain on CLASS Score	13
4.3 Pre- and Post- CLASS Score	15
4.4 CLASS Categories	16
4.5 The Effect of High-School Experience	18
4.6 Direct Correlation	20
4.7 Specific Considerations in CLASS	21
5. Conclusions and Future Work	28
-Acknowledgements	29
-Sources	29
-Appendix A: CLASS Questions	30
-Appendix B: CLASS Categories	31
-Appendix C: Interview Problem Solving Questions	31

1. Introduction

The purpose of breadth courses are “to provide students with meaningful exposure to areas of knowledge outside their primary area of study.”, and to encourage students “to explore new and different analytical and critical methods”, in order to give students a well-rounded education and foster interdisciplinary thinking.¹ The University of Toronto (UofT) requires Arts & Science students to choose courses from at least four of five different categories, three in arts and two in science, one of which is ‘The physical and Mathematical Universes’:

“The Physical and Mathematical Universes: an understanding of theories of the physical world and mathematical models, and the ability to apply them and to make and evaluate observations relating to them.” “Courses in the physical sciences study the constituents and mechanisms that govern the natural world, with an emphasis on non-living systems. Such courses develop an understanding of scientific methodology (the formulation of hypotheses, experimental design, and quantitative analysis methods), its application and its limits.”¹

Two courses within this category are PHY100 – ‘The Magic of Physics’ - and PHY205 – ‘Physics of Everyday Life’. PHY100 and PHY205 are aimed at non-science students who need to satisfy their breadth requirements and who (hopefully) have an interest in learning some basic physics. PHY100 explores a large range of physics topics, ranging from Newton’s laws to quantum mechanics. The calculations in this course are kept simple, but students choose topics to explore in depth on two written assignments. PHY205 covers fewer topics in greater detail, with more emphasis on quantitative methods and less on written expression. As the course titles suggest, PHY205 focuses on physics that students encounter in their non-academic lives, such as kinematics, forces, fluids, heat, and waves. Both courses are taught in a traditional format, with material being introduced through lectures and then discussed during weekly tutorials. PHY205 also actively involves students during lectures by making use of clickers. Tutorials for this course have students doing experiments that build upon concepts they were taught in class. The tutorials for PHY100 were based on class discussion and group problem solving.

While the mastering of concepts in these courses is evaluated through examination, this does not necessarily guarantee that the aims of breadth courses set out by the curriculum committee are being met. This report evaluates the effect of physics breadth courses at UofT on students’ attitudes towards physics, using the Colorado Learning Attitudes About Science Survey (CLASS).

CLASS consists of 42 statements to which students are asked to respond on a five point scale.² An over-arching theme of CLASS is discerning whether a student looks at physics as a coherent body of knowledge (the expert view), or as a collection of unrelated facts and formulas (the novice view). A full list of the statements can be found in appendix A. In regards to the goals of breadth courses, questions 3, 28, and 30 refer to interdisciplinary thinking; questions 8, 17, 22, 26, 39, and 42 evaluate whether the student has absorbed physics methodology; questions 27 and 41 address the nature of science as a method of inquiry and its role in broader society.

2. CLASS Literature Review

2.1 The Development of CLASS v III

CLASS was developed with the specific goal of probing beliefs about the process of learning physics.² The survey in its current form (3rd edition) contains 42 statements, that students are asked to respond to on a five point scale from A-‘strongly disagree’ to E – ‘strongly agree’. The expert answer for each statement was determined by consensus from the answers given by physicists. Out of the 42 statements, 36 elicited strong consensus, so these 36 are scored. Question 31 is a qualifier that requires students to answer ‘D’ in order to disqualify data from students who improperly fill out the survey. Care was taken to ensure that the statements on CLASS were written clearly and could be applied to many different types and levels of physics courses. CLASS contains eight empirically determined, non-mutually-exclusive categories.² Studies report scores for the individual categories as well as the overall score.^{2,3,4,5,6}

CLASS started as a collection of statements taken from the surveys MPEX and VASS. Each was modified by reading it to students and asking students to repeat it back in their own words. This ensures that the vocabulary used is familiar to students. Students seem to use the word physics in three different ways²: 1) to describe their course, 2) to describe the field of study, and 3) to describe physics in the real world. It is the third notion of physics to which CLASS aims to refer.

CLASS underwent a series of validation steps in its development:

- 1) The statements were reviewed by 16 physics faculty members to determine the expert responses. The statements were adjusted if faculty members thought they were ambiguous.
- 2) Statements were reviewed by 34 students from 6 different physics courses to ensure clarity.
- 3) Preliminary categories were developed using surveys from several thousand students.
- 4) Validity was tested by correlating pre-course beliefs and grades.
- 5) Validity was tested to make sure the results made sense. For example, it is expected that physics majors score better than other students, particularly non-science students.

These steps resulted in revisions and each augmented version underwent the above steps, leading to the product used today, CLASS v. III. In step two, each student first took the survey and were asked for demographic information. Then the interviewer read each statement to the students and asked the student to talk about their answer and interpretation of the statement.² Some statements on the latest version of CLASS resulted from revelations about student attitudes garnered from the interviews, such as how students connect physics with their daily lives.

Categories were developed through a mix of two techniques – *statistical analysis* and *predeterminism*. The statistical method puts no constraints on the original survey, which is taken by a large sample of students. Their answers provide the categories via factor analysis, which groups statements together based on correlations of student responses. Grouping the statements gives rise to a set of factors that give an oblique basis set that most closely spans the space of students’ answers.² A benefit of this method is the guarantee of having statistically sound

categories. In addition, it provides insight into the way that students think, which might have been previously unknown to educators. For example, factor analysis on preliminary CLASS responses showed that students link sense-making with effort.² This was verified in the interviews where students expressed that sense-making, as opposed to rote memorization, is an extra effort they put in when they think it will pay off. Factor analysis also revealed that there are two factors in how students relate physics to the real world: there is the question of whether they believe physics from the classroom translates into important concepts in the real world; and there is the question of whether or not they care.² The drawbacks to this method are that not all categories will necessarily be useful for analysis, and that categories are mutually-exclusive, which excludes statements from certain categories where they might be appropriate.

Predeterminism starts with the assignment of categories by the survey creators, which produces categories that reflect their beliefs about useful categorization of aspects of learning.² A benefit to this method is the guarantee of useful categories that address what the creator is interested in probing. However, this does not necessarily produce statistically valid categories, nor does it give any new insight into the way students think about learning physics. CLASS was developed using a combination of these two methods, so as to ensure categories that are both empirically valid and useful for educators.² First, statistical analysis was used on predetermined categories borrowed from other surveys, grouping the statements into new categories. Statements were not restricted to only one category and not all statements were forced into a category. Further statistical analysis was done on the categories using only the set of statements that might possibly be categorized. This process is shown in figure 1.

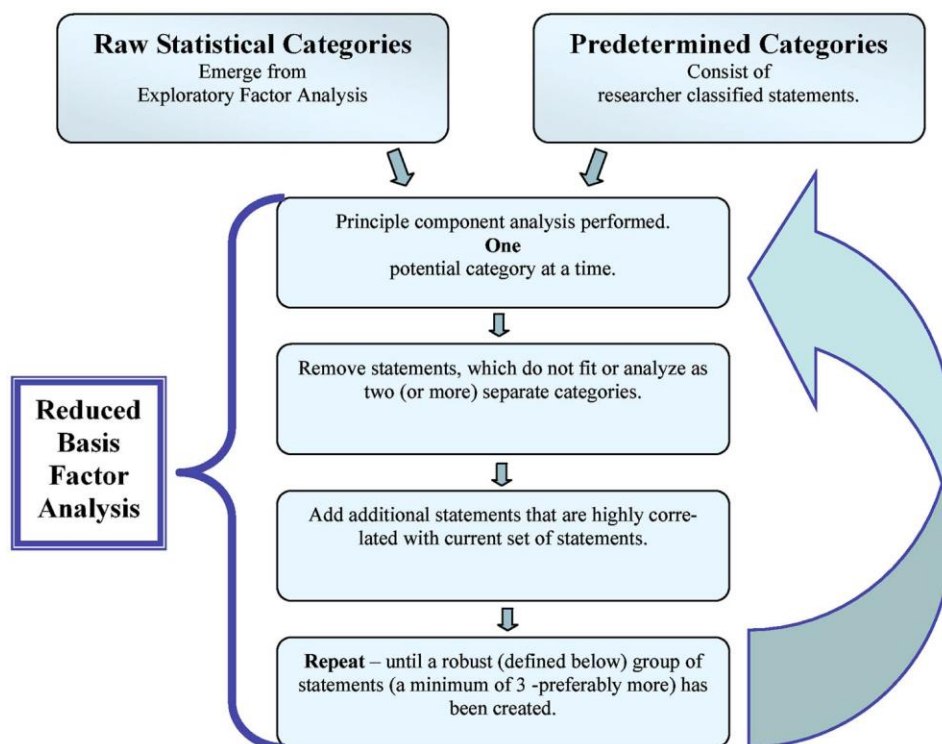


Figure 1: A Flowchart of the development process for CLASS²

The data were taken from three courses: a calculus-based course, an algebra-based course, and a physics-for-non-scientists course. The data were put through several cycles of this analysis, each time adding or subtracting statements from categories, in order to determine the most statistically sound and useful categories. The eight categories were named after-the-fact.

Scoring is done on a three point scale, as opposed to the five point scale that answers are collected on.^{2,3,4,5,6} This amounts to combining strongly agree with agree and combining strongly disagree with disagree. The reason for this is that students perceive the difference between agree and strongly agree differently, so there is nothing close to a consensus on the relative strength of these answers. However, it is still important to include both these options in the survey because it can help to coax non-neutral answers out of students. In interviews, students revealed a number of reasons for responding neutrally: 1) Did not know how to answer, 2) did not have an opinion, 3) conflicting opinions based on various experiences with physics, 4) conflict between their own opinions and what they perceive the correct answer to be. These varied reasons for neutrality led the CLASS creators to recommend that scoring be done by calculating percent-expert and percent-novice scores, while ignoring neutral responses.² However, most papers only report percent-expert scores.^{3,4,5,6}

2.2 Uses and Results of the Administration of CLASS

Between fall 2003 and fall 2005, the creators of CLASS administered the survey to over 7000 students in 60 different physics courses. Most instructors found a drop in expert-like views over the term.^{2,3}

Table 1: Percent-expert scores from a large-state research university calculus-based introductory physics course (N = 397).² This is representative of typical CLASS results.

Category	Pre	Standard deviation	Post	Standard deviation
Overall	65(1)%	16	59(1)%	20
Real world connections	72(1)%	28	65(2)%	32
Personal interest	67(1)%	28	56(2)%	32
Sense making/effort	73(1)%	22	63(1)%	27
Conceptual connections	63(1)%	25	55(1)%	28
Applied conceptual understanding	53(1)%	25	47(1)%	28
Problem solving general	71(1)%	23	58(1)%	28
Problem solving confidence	73(1)%	27	58(2)%	33
Problem solving sophistication	61(1)%	29	46(2)%	32

Gray et al looked into students' perceptions of how physicists would respond to CLASS.³ Students were asked to indicate both their own answer and the answer they would expect from a physicist on both the Force Concept Inventory (FCI) and CLASS. Three courses were used: a calculus-based course, an algebra-based course, and a non-science-majors course. Results show that although the students scored around 50-60% on their own views with a decline over the term (in line with other studies), they were very adept at guessing the expert-like responses to the

statements, scoring around 80%. Sixty-six American physicists were also asked to take the CLASS for this study. Their average score was 91.4%.³

Table 2: Questions for which students did not successfully guess the expert response³

Statement (expert response)	Faculty ^b (percentage favorable)	Pre-“physicist”			Post-“physicist”		
		Phys I-Calc (%)	Phys I-Alg (%)	Phys of Sound (%)	Phys I-Calc (%)	Phys I-Alg (%)	Phys of Sound (%)
8. When I solve a physics problem, I locate an equation that uses the variables given in the problem and plug in the values. (Disagree)	83	27	27	21	31	41	29
12. I cannot learn physics if the teacher does not explain things well in class. (Disagree)	63 ^c	27	22	33	32	28	28
18. There could be two different correct values for the answer to a physics problem if I use two different approaches. (Disagree)	79	63	50	41	68	59	50
27. It is important for the government to approve new scientific ideas before they can be widely accepted. (Disagree)	100	56	49	55	58	47	57
22. If I want to apply a method used for solving one physics problem to another problem, the problems must involve very similar situations. (Disagree)	91	59	50	52	60	49	49

The original CLASS interviews were used to provide possible explanations as to why students show such a split between their own views and their projective views: Students tend to perceive the questions they are required to answer in a classroom setting as less authentic than the real life problems physicists work on. They see physics research as having less memorization and repetition as students’ homework.²

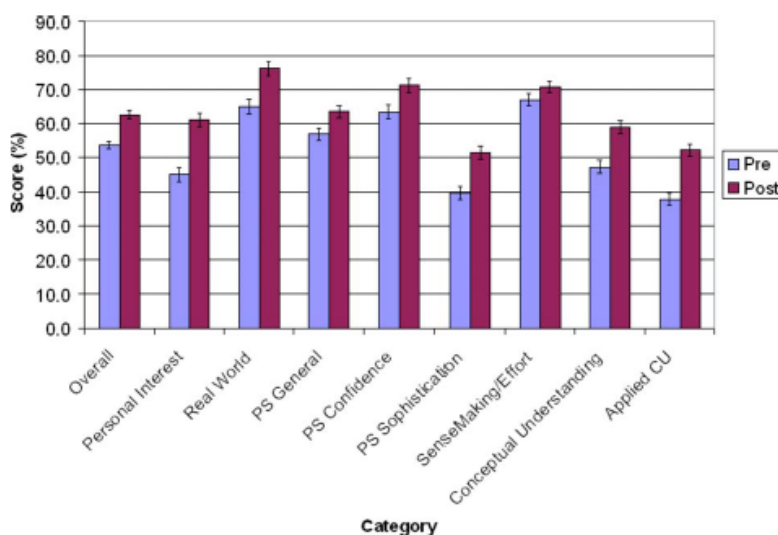
The majority of physics courses taught in Canada and the United States are largely traditional, in that they are mostly lecture-based. CLASS has been used to investigate the effects of other curricula on student beliefs about physics. Courses in which instructors explicitly address beliefs about learning physics see better results than other courses.² One curriculum that does this is ‘Physics and Everyday Thinking’ (PET), which has been shown to have a positive impact on conceptual learning, as tested through the FCI.⁴

PET was specifically designed for the physics education of elementary teaching candidates. Many studies suggest that teachers, and therefore their students, do not hold an expert-like view of the nature of science; where the correct nature of science is defined as “the use of empirical standards, logical arguments, and skepticism.”⁴ The goal of PET is to improve these views during teacher candidates’ science education through the explicit instruction of this issue. It focuses on interactions, energy, forces, and fields.

Otero and Gray administered CLASS to 360 students over nine PET courses at seven universities. The individual courses had between 9 and 100 students enrolled. The average of all of the courses was a gain of 8.8% expert-like. The results for each institution and in individual categories are summarized in Table 3 and Figure 2, respectively.

Table 3: Overall Percent-Expert Scores From Nine PET Courses.⁴

Course	Type	Enrollment	Curriculum	Pretest	Posttest	Shift
1	Community college	13	PET	69.9 (2.6)	73.8 (4.2)	3.9 (3.6)
2	Community college	30	PET	53.6 (4.3)	67.0 (2.6)	13.3 (3.1) ^a
3	Research university	100	PET	51.6 (2.0)	58.3 (2.2)	6.7 (2.0) ^a
4	Regional university	32	PSET	49.5 (3.0)	59.0 (3.7)	9.5 (3.6)
5	Regional university	32	PSET	49.8 (3.7)	59.6 (3.8)	9.8 (2.8)
6	Regional university	25	PET	51.8 (2.7)	68.3 (3.3)	16.5 (3.8) ^a
7	Technical university	30	PET	51.6 (4.5)	62.1 (3.7)	10.6 (4.3)
8	Regional university	48	PET	64.2 (3.0)	70.4 (3.8)	6.3 (2.5)
9	Regional university	50	PET	55.0 (3.6)	63.1 (4.0)	8.1 (2.7) ^a
All students				53.8 (1.15)	62.6 (1.18)	8.8 (1.1) ^a

Figure 2: Percent-expert scores in each category, averaged from nine PET courses⁴

Most curricula omit explicit discussion of epistemology because of time constraints, but the Physic by Inquiry (PbI) curriculum covers epistemological concerns implicitly.⁵ PbI was designed at the University of Washington and is aimed at teaching candidates and non-science majors. The course is divided into various topics (kinematics, light, astronomy . . .), each containing a number of questions. The students work in small groups (2-3) and perform experiments in order to answer the questions and build up physical models based on their own results. This is supposed to reflect the way in which scientists generate data, although this is not explicitly explained to the students. There are no lectures and very few large group discussions. Students keep a lab notebook where they record their results and develop models. These notebooks can be used in exams, which are based on application of familiar concepts to unfamiliar situations. Many studies show that PbI is effective in teaching physical concepts to students as measured by the FCI.⁵

Lindsey et al found that students in PbI courses showed statistically significant positive shifts on CLASS that are comparable to the shifts seen with PET courses.⁵ In particular, there were large shifts in the problem solving categories. Students from five different institutions filled out CLASS at the beginning and end of term.⁵ In addition to the survey, some students were asked for written responses to some follow-up prompts for the problem solving questions. Overall the results show significant positive shifts, although these shifts differ between institutions and between semesters at the same institution. The pre-course values range from around 50% to 60%⁵, in line with other studies.^{2,3,4,6} As shown in figure 3, gains ranged from 1.9% to 25.1%, whereas PET courses have produced gains of 3.9% to 16.5%, so PbI appears to measure up to PET in terms of improving students' attitudes.⁵

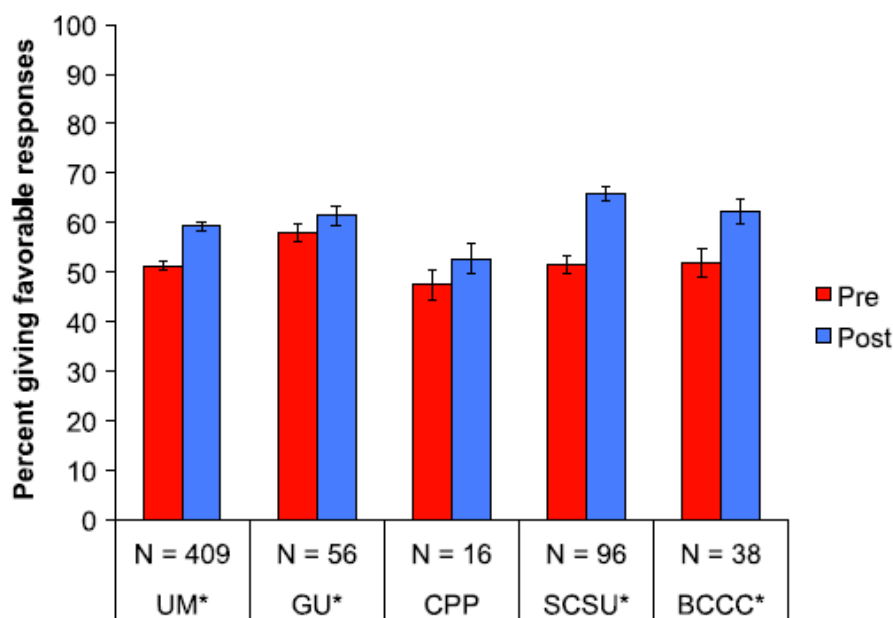


Figure 3: Average CLASS scores in five institutions employing PbI.⁵

This study also suggests that PbI helps combat the ‘Plug and Chug’ mentality embodied in CLASS statement 8. They asked students to discuss their attitudes towards problem solving and received responses including the following:

“I have learned that there is more meaning behind physics equations that just formulas for you to plug numbers into (for example, proportional reasoning is more than just cross-multiplying).”⁵

“Physics equations, while they can definitely be used for calculations, exist to explain the relationship between variables. In this class, we very rarely had to simply plug values into an equation and allow the calculator to find the answer for us. More frequently, we were expected to recognize how certain variables related to or affected each other.”⁵

Here is a comparison of one student’s views from before and after the course:

Before - “To solve a problem you usually need to know what rule or equation to use and then you need to plug the numbers from the problem into the equation to calculate an answer.”⁵

After – “I used to think that math and science problems (involving math) were only to solve the problem and find the numerical answer. I now am able to apply the answer, not only to the problem, but also to everyday life”. . . “I can understand what I am solving and why.”⁵

These studies and many more support the idea that some less traditional approaches to teaching physics have a positive effect on both student attitudes and on conceptual understanding. However, many institutions want to improve students’ attitudes towards physics, but do not want to veer away from traditional pedagogy. Of particular interest to physics lecturers at UofT are the CLASS results from Ryerson University. Ryerson employs a fairly traditional physics curriculum and has found positive shifts in its students’ CLASS scores.⁶

At Ryerson, students were given the FCI and CLASS before and after completing an introductory calculus-based physics course. The course included two lectures and one tutorial per week. Both clickers and in-class demonstrations were used, similarly to PHY205. Students completed the surveys voluntarily and anonymously, with no incentives to complete CLASS. 155 surveys were used in the data.⁶

Students at Ryerson showed a positive gain overall and in most of the categories. The exceptions were Problem Solving General and Sense Making and Effort. They compared their results to similar courses at the University of British Columbia (UBC) and at the University of Colorado (UC), both of which found the typical drop in CLASS scores. The lowest scored categories at all three universities were *applied conceptual understanding* and *problem solving sophistication*.⁶ The full table of results from this study will be shown later in this paper so as to compare their results with ours.

3. Methods

CLASS was administered in January 2013, in the first week of PHY205 and in the second week of PHY100 to obtain the pre-course scores. Students in both courses were surveyed again in the last week of the course (April) to obtain the post-course scores. The surveys were written in tutorials with no incentive offered for completion. The students were asked to indicate their fields of study from ten options: *physical science or mathematics, computer science, life science or kinesiology, engineering, social science, philosophy, humanities, fine arts, commerce, or other*. Upon analyzing the data, it was determined that students studying Social Science, Philosophy, and Humanities were very similar. They were therefore grouped together as ‘liberal arts’. Because there were so few fine arts students, they were also lumped in with this group which we now call ‘arts’. The commerce students behaved distinctly, and so were kept as a distinct grouping, as were ‘life science’ and ‘computer science’. Only three students identified as engineering, so there were placed with ‘physical science’. This produced five categories: physical science, computer science, life science, commerce, and arts

Only surveys with the correct directed answer D for question 31 were considered. To ensure that the same group of students was being considered throughout the study, only students who wrote both surveys were used for the data. Any students who reported their major as 'Other', or who reported conflicting majors in parts A and B of the demographic question were discarded.

The CLASS scores were calculated using the 36 score-able questions. Although the students answered the questions on a five point scale, the survey was marked on a three point scale, in line with other usages of CLASS. CLASS studies calculate both the percent-expert scores and the percent-novice scores; however, most papers end up only reporting the expert score, because also including the novice score becomes cumbersome. In our case, leaving out the novice score seemed to be neglecting important data by not distinguishing between novice and neutral answers. Therefore, the scores reported here are the differences between percent-expert and percent-novice scores:

$$\text{presented score} = \text{actual expert score}/0.36 - \text{actual novice score}/0.36$$

This is the case with all of these results, including scores on individual categories and specific questions. A score of 50% really means that 50% more students were expert-like than novice-like. A score of -50% really means that 50% more students were novice-like than expert-like. The exception is table 4 which compares UofT students with those from other universities. These scores are percent-expert only.

The standard error for average scores was calculated assuming a normal distribution.

$$SE = \frac{\sigma}{\sqrt{n}}$$

The error for individual questions assumed a binomial distribution for each percent-expert and percent-novice.

$$E_{\text{bin}} = \sqrt{\frac{p*q}{n}}$$

These errors were combined for the error of the difference.

$$Error = \sqrt{E1^2 + E2^2}$$

Similarly, the errors for the averages of gain were calculated by combining the errors of pre- and post-course scores.

$$Error = \sqrt{E_{pre}^2 + E_{post}^2}$$

Some variables were directly correlated by plotting them against each other and obtaining the correlation coefficient (calculated by Excel) and its error.

$$SE = \sqrt{\frac{1-r^2}{n-2}}$$

All averages and errors were taken to one significant figure at the end of the calculation.

Some variables were studied over all of the students from both courses, as opposed to just within each course. This was done with incoming students, who are not expected to differ between courses before the term, and with any data that showed similar trends in both courses.

Sixty-one surveys were used from PHY100, of which 61% were from arts students, 16% commerce, and 23% science. PHY205 had considerably more science and commerce students, with 26% arts, 36% commerce, 16% physical science, 12% life science, and 10% computer science. One-hundred and eighty-four surveys were used from PHY205, for a total of two-hundred and forty-five surveys collectively.

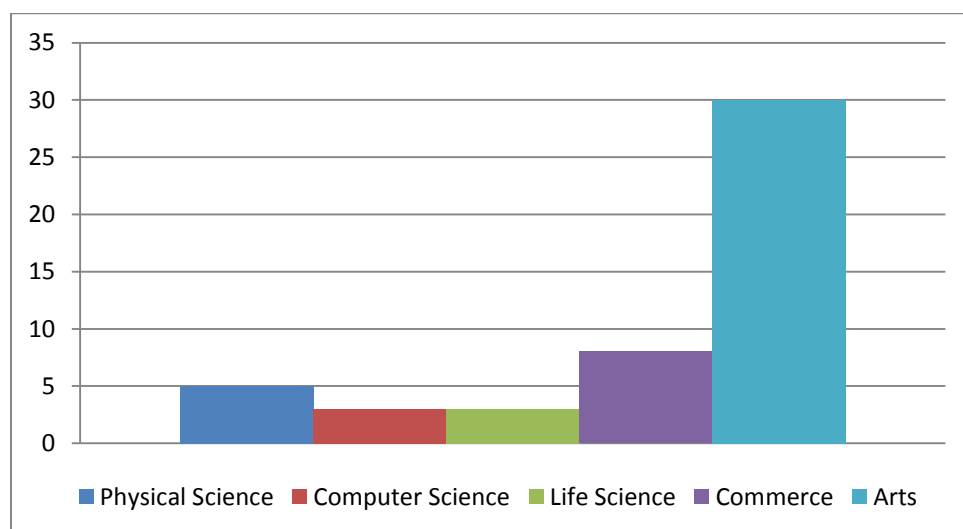


Figure 4: PHY100 student respondents by program of study

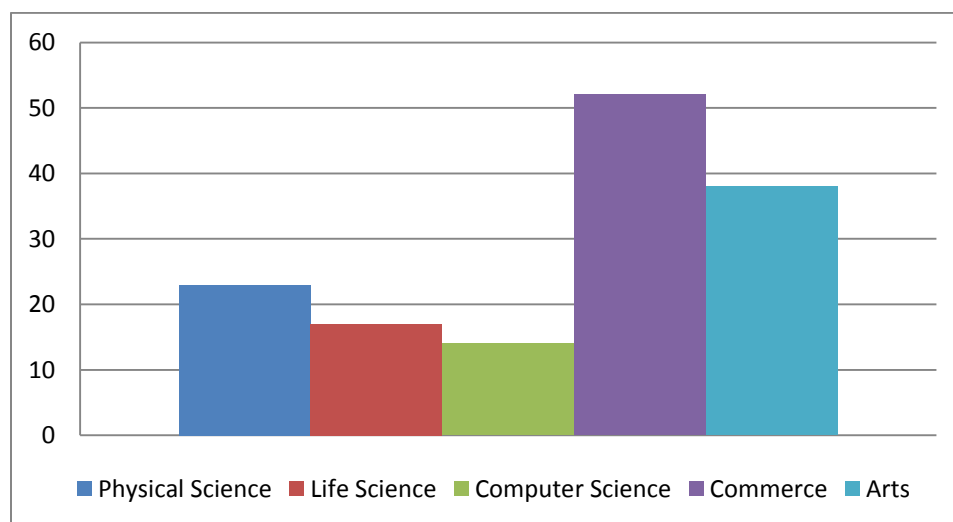


Figure 5: PHY205 student respondents by program of study

In PHY100, 50.6% of students had no high-school physics, 16.1% had grade eleven only, and 33.3% had both grades eleven and twelve. In PHY205, 28.2% of students had no high-school physics, 25.1% had grade eleven only, and 46.7% had both grades eleven and twelve.

All of the results are based only on the students whose surveys met the qualifications stated above. Many of these results can be taken as representative of the greater population, but not necessarily all of them.

4. Results/Discussion

4.1 Final Course Grades

The average grades in this study are somewhat higher than the actual course averages. This could be because students included in this study must have completed the post-course survey, and therefore must have attended the last tutorial. It is possible that students who attended the final tutorial were more likely to be conscientious students and earn higher grades.

The average overall grade for the surveyed PHY100 students was $73.7\% \pm 1.9\%$ (Fig 6). The five physical science students did not do particularly well. Both the arts and commerce students performed decently, with a statistically insignificant difference between the two groups' performances. The average grade in PHY205 was $77.9\% \pm 0.9\%$ (Fig 7). The physical science students performed better in PHY205 than in PHY100, but were still below the course average. Arts students had the lowest average, $73.9\% \pm 1.7\%$, life science students the highest ($83\% \pm 1.9\%$). Of the surveyed PHY205 students, only four failed the course. This is much less than the total number of failures in the course, indicating that information from the surveyed sample of students cannot fully represent the total population.

These grade distributions might be due to the quantitative nature of assessment in PHY205, compared to the focus on writing in PHY100. Arts students, for whom these courses are largely designed, seem to fare better in PHY100, which has a large written component. When interviewed, arts students claimed that writing was one of their academic strengths, whereas they had mixed perceptions of their mathematical abilities. The few physical science students did not do well in either course, but such students must be atypical since students who have taken regular first year physics courses are excluded from taking either PHY100 or 205.

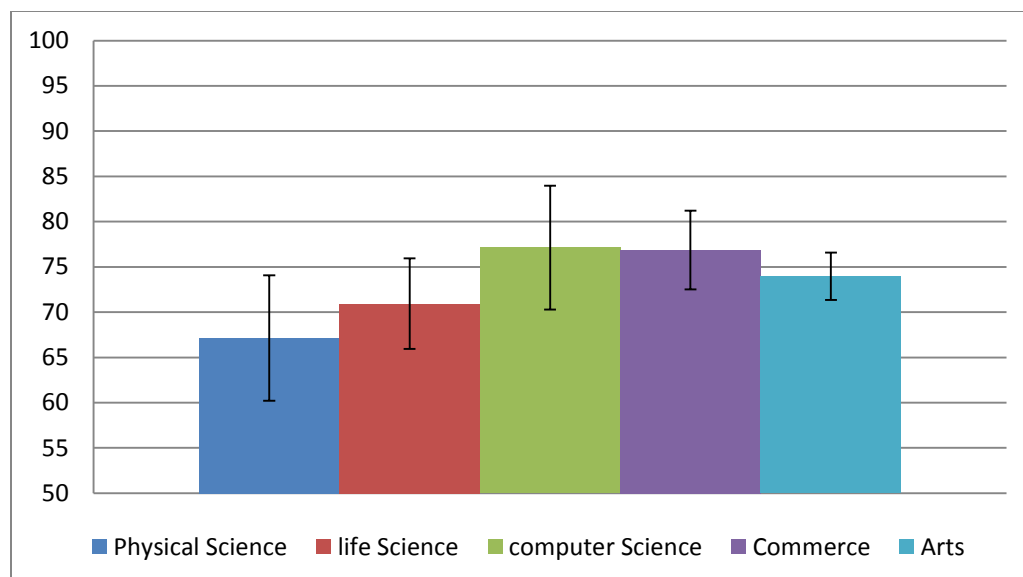


Figure 6: PHY100 final course grades by program of study

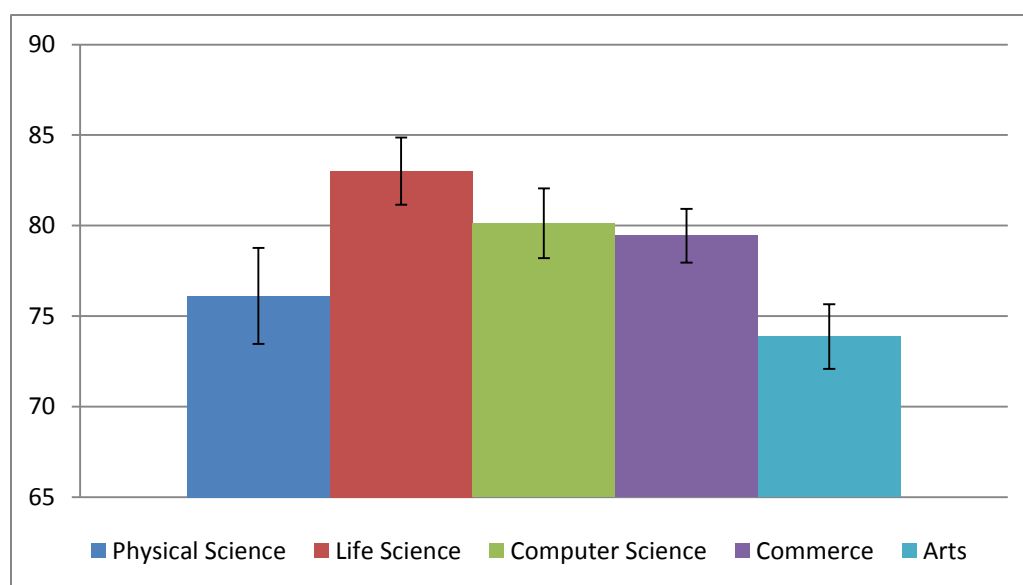


Figure 7: PHY205 final course grades by program of study

4.2 Gain on CLASS Score

The percent gain is defined as the difference between the pre-course and post-course CLASS scores. The average gain was $0.9\% \pm 5.1\%$ for PHY100 (Fig 8) and $4.1\% \pm 2.5\%$ for PHY205 (Fig 9). Both courses had an almost neutral effect, individually and combined (Fig 10).

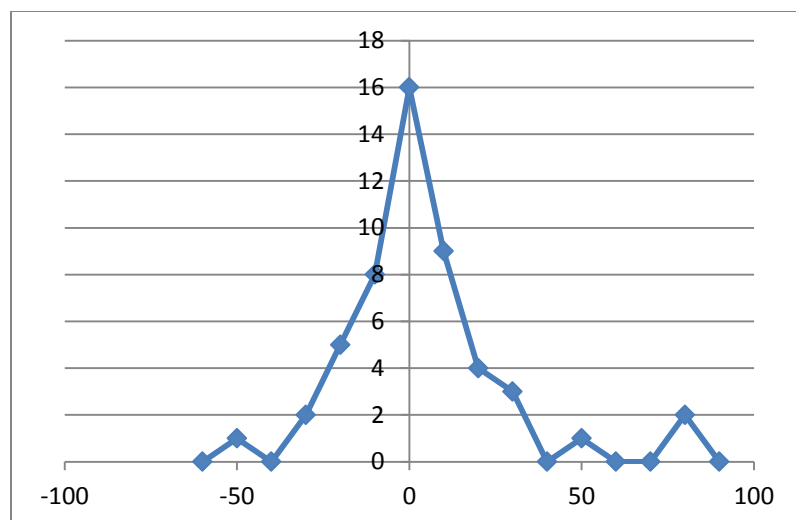


Figure 8: Histogram of percent gain on CLASS scores in PHY100

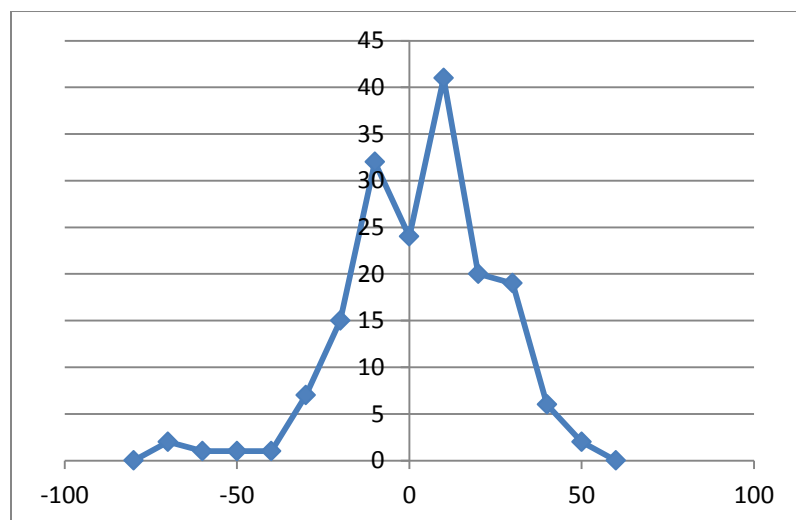


Figure 9: Histogram of percent gain on CLASS scores in PHY205

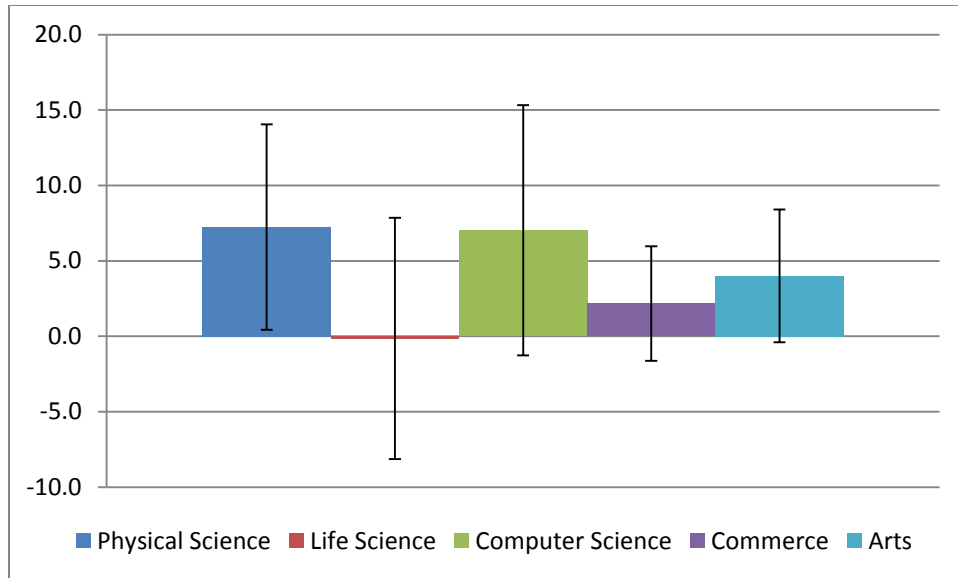


Figure 10: Percent gain on CLASS scores in both courses by program of study

4.3 Pre- and Post-CLASS Scores

Combining the courses, the average pre-course score was $39.0\% \pm 1.5\%$ (Fig 11), and the average post-course score was $42.2\% \pm 1.5\%$ (Fig 12). The pre-course scores were 2.8σ better for commerce students than for arts students. The post-course scores were 2.4σ better for commerce students than for arts students.

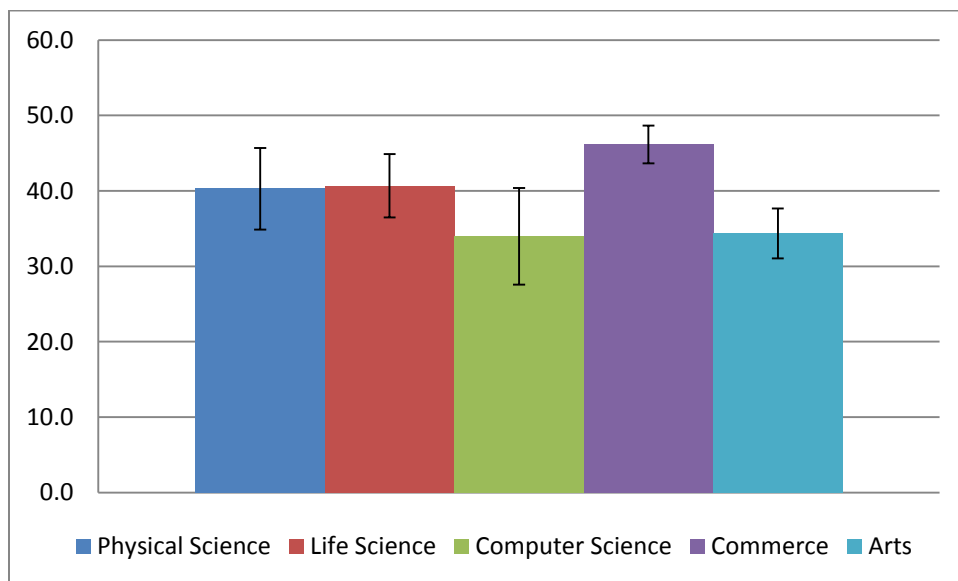


Figure 11: Pre-course CLASS scores in both courses by program of study

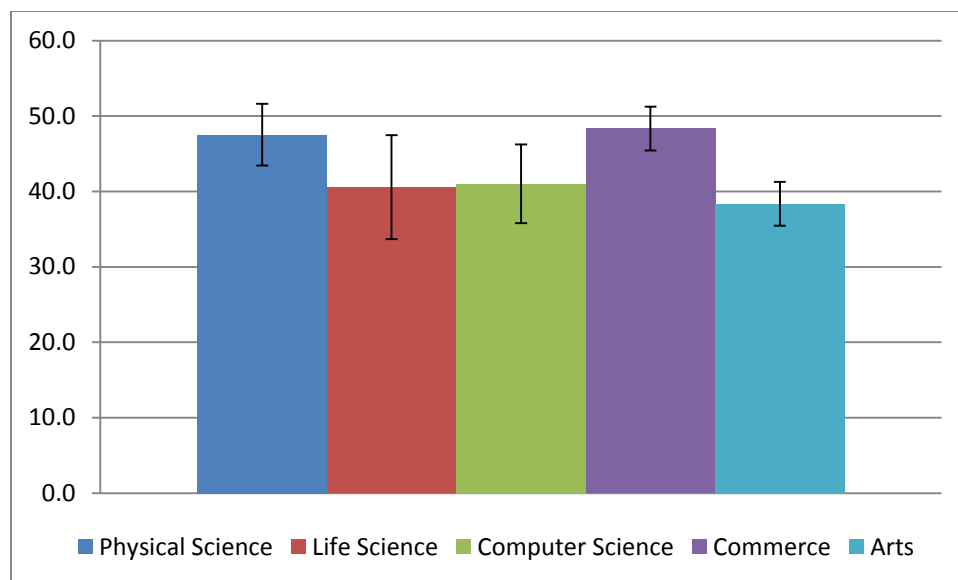


Figure 12: Post-course CLASS scores in both courses by program of study

4.4 Categories of CLASS

The eight categories in CLASS are *real world connections* (RWC), *personal interest* (PI), *sense making/effort* (SME), *conceptual connections* (CC), *applied conceptual understanding* (ACU), *real world connections* (RWC), *problem solving general* (PSG), *problem solving confidence* (PSC), and *problem solving sophistication* (PSS). For incoming students in both courses the best scoring categories were RWC ($57.4\% \pm 2.5\%$) and SME ($60.7\% \pm 2.2\%$), then PSG ($43\% \pm 2.2\%$), PSC ($46.6\% \pm 3\%$) and PI ($47.2\% \pm 2.6\%$), then CC ($35\% \pm 2.7\%$), then PSS ($24.2\% \pm 3\%$) and ACU ($10.6\% \pm 2.4\%$). The ordering of the categories (Fig 13) stayed nearly the same after the term. The high scores in RWC and PI may be due to the fact that students in these courses chose to take physics instead of another breadth course; therefore they probably have more interest in physics than the general student population.

The questions in ACU and PSS involve the application of conceptual understanding to problem solving. This coupling of skills is more difficult for students than dealing with conceptual understanding and problem solving separately. Students often memorize a physical law, but fail to apply it correctly in different contexts: In interviews, students had much more trouble solving conceptual questions than performing calculations. They all successfully answered interview questions (Appendix C) one and three, but had difficulty on two and four. When asked about paddling a canoe in interview Q2, they all responded with “because of Newton’s Third Law”, but had difficulty explaining how the law applied to this situation. In the summer term of PHY100, students were asked if an electron or a proton would fall faster in a gravitational field. Some of them recited *Galileo’s Law of Falling Bodies*, then proceeded to write that the proton would accelerate faster, demonstrating a failure to properly apply conceptual understanding.

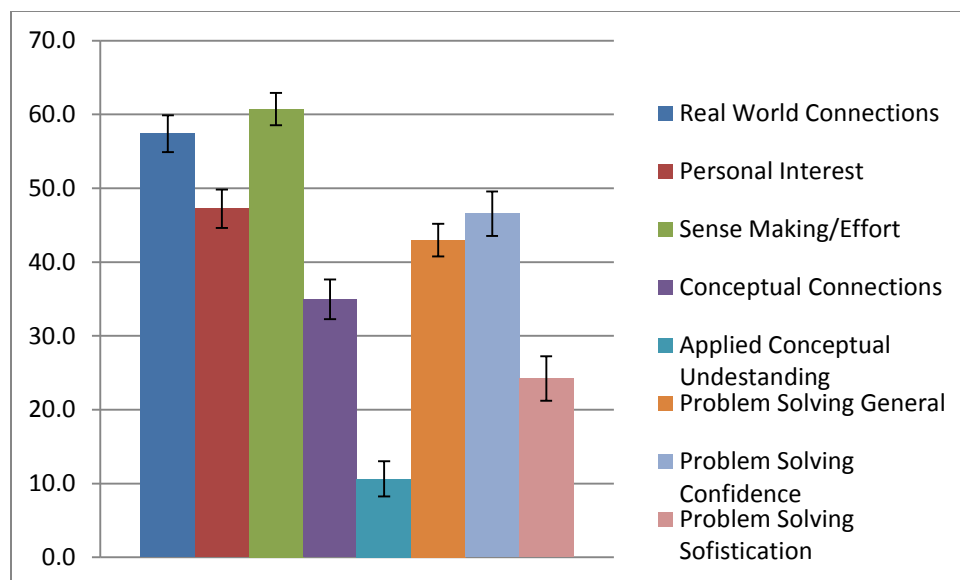


Figure 13: Pre-course CLASS scores in both courses by category

There were no significant gains or losses in PHY100 (Fig 14). In PHY205 there were significant gains in PI ($6.3\% \pm 3.9\%$), PSC ($10.3\% \pm 4.6\%$), and PSS ($6.7\% \pm 4.5\%$) (Fig 15). It is not surprising to see a gain in PSC and PSS because the course assignments gave students problem solving experience, building their skill and confidence.

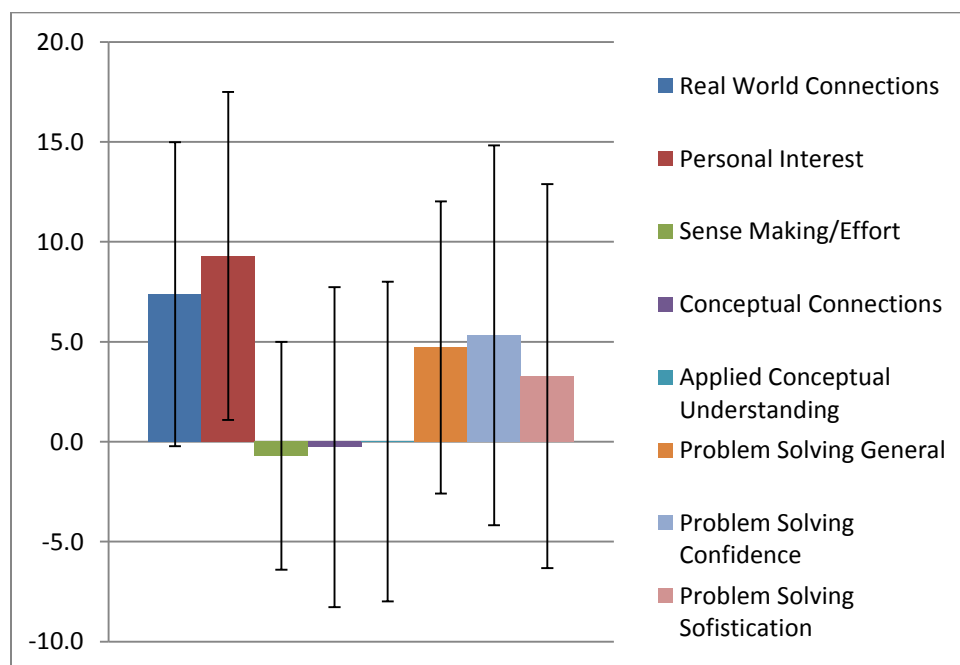


Figure 14: PHY100 percent gain on CLASS scores by category

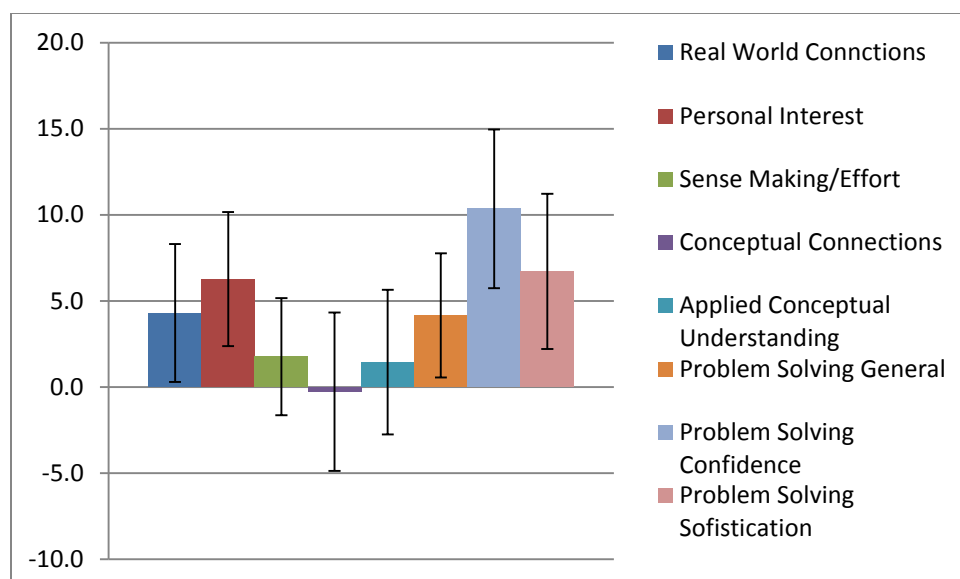


Figure 15: PHY205 percent gain on CLASS scores by category

The table below shows a comparison between students at UofT and students at UC⁶, UBC⁶, and Ryerson⁶. Unlike the other results presented in this paper, these scores are based on percent-expert only. At all of the universities, the lowest scoring categories were ACU and PSS, while SME was at or near the top.

Table 4: A comparison of CLASS scores between four institutions. Adapted from *Attitudes about science and conceptual physics learning in university introductory physics courses*.⁶

Institution	UC			UBC			Ryerson			UofT PHY100			UofT PHY205		
	pre	post	gain	pre	post	gain	pre	post	gain	pre	post	gain	pre	post	gain
Overall	65	59	-6(2)	48	46	-2(4)	55.8	57.7	1.9(2)	56.3	57.8	1.5(3)	58.7	60.9	2.2(1.6)
PI	67	56	-11(3)	40	36	-4(6)	55.4	57.3	1.9(4)	57.1	63.1	6(5)	61.9	64.2	2.3(2.7)
RWC	72	65	-7(3)	50	46	-4(8)	64.5	69.7	5.2(5)	70.1	72.5	2.5(5)	68.5	72.2	3.6(2.8)
PSG	71	58	-13(2)	50	45	-5(6)	58.3	61	2.7(2)	59.6	61.5	1.8(4.6)	65.5	57.3	-8.3(2.1)
PSC	73	58	-15(3)	51	45	-6(7)	56.9	60.8	3.9(2)	58.6	60.2	1.6(5.7)	61.6	68.3	6.7(3.2)
PSS	61	46	-15(3)	34	28	-6(6)	39.9	42.7	2.8(2)	44.0	45.1	1.1(5.5)	48.4	53.3	4.9(2.8)
SME	73	63	-10(2)	63	56	-7(6)	69.5	67.8	-1.7(2)	67.7	67.0	-0.7(3.9)	71.2	71.9	0.7(2.2)
CC	63	55	-8(2)	43	41	-2(6)	49.6	51.7	2.1(2)	53.0	53.6	0.5(4.8)	57	58.1	1.1(2.8)
ACU	53	47	-6(2)	34	32	-2(4)	38.7	42.3	3.6(2)	40.3	39.8	-0.5(4.2)	43.4	45.5	2.1(2.5)

4.5 The Effect of High-School Experience

More physics experience in high-school seemed to correspond with slightly higher grades in PHY205 (Fig 17), but not in PHY100 (Fig 16). This is perhaps due to students with high-school physics being more comfortable with the algebraic skills needed for PHY205. Students in both courses who had completed grade twelve physics scored higher on CLASS at the start and end of term (Fig 18&19), but did not gain more over the term. The advantage of completing high-school

physics seen in this study is much smaller than what has been seen in other studies using CLASS.^{2,6}

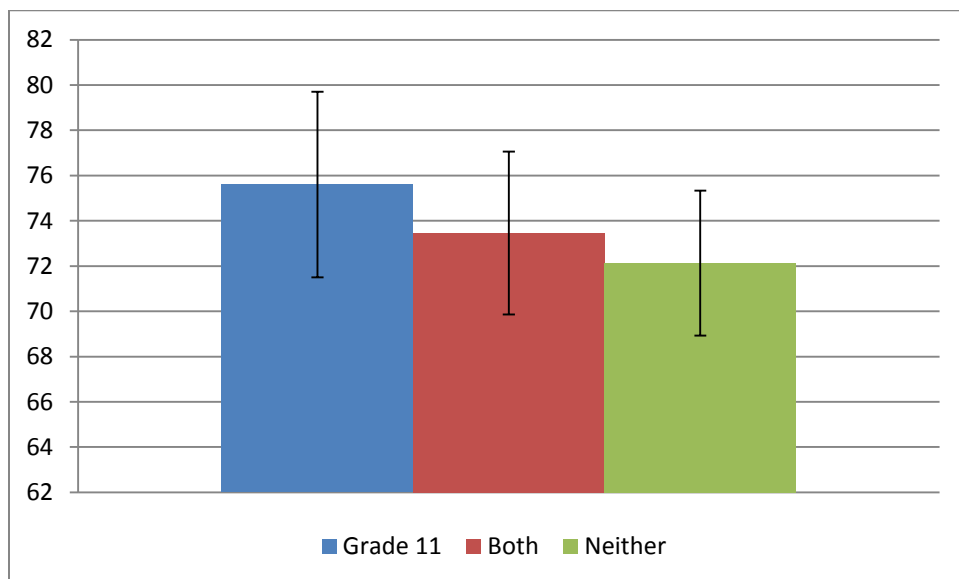


Figure 16: PHY100 final grades by level of high-school completed

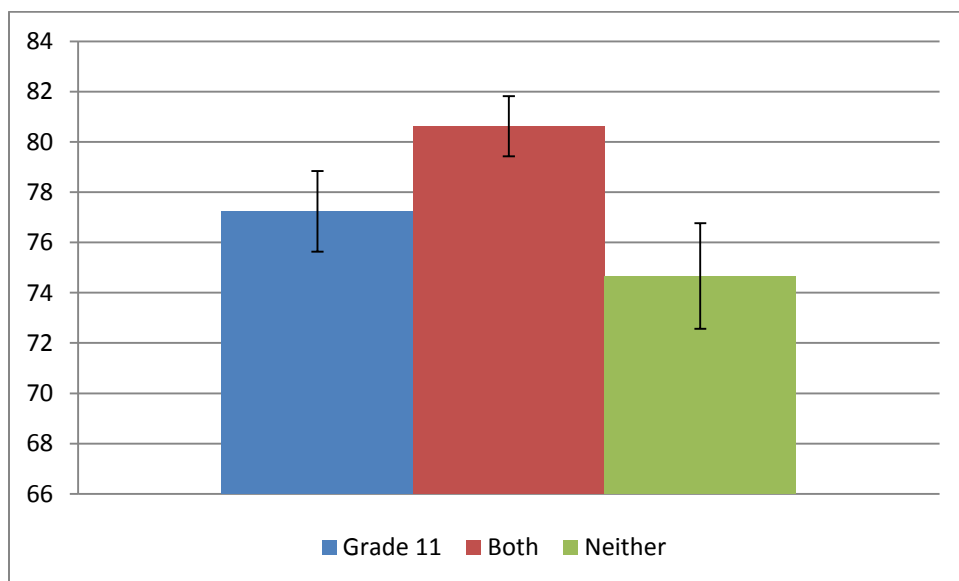


Figure 17: PHY205 final grades by level of high-school completed

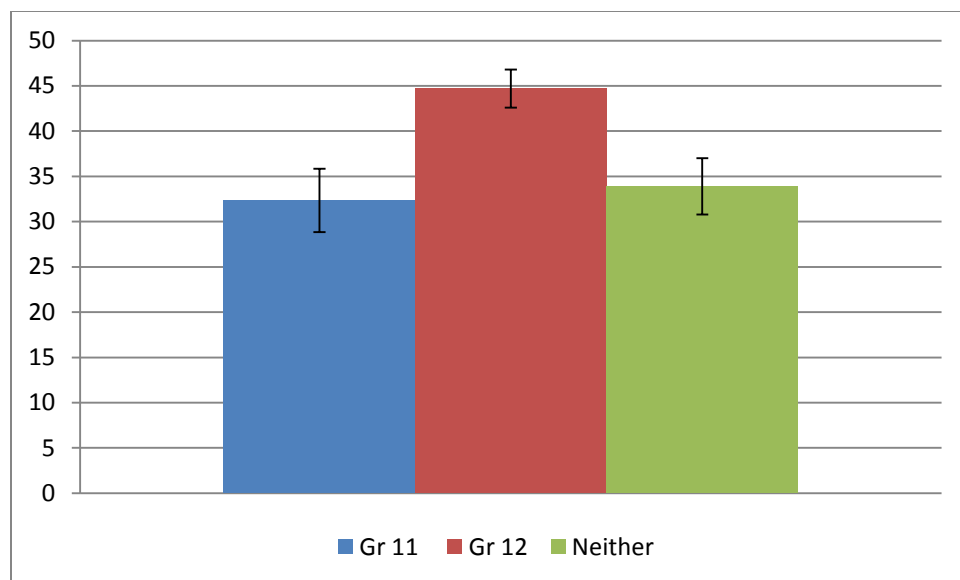


Figure 18: Pre-course CLASS scores in both courses by level of high-school completed

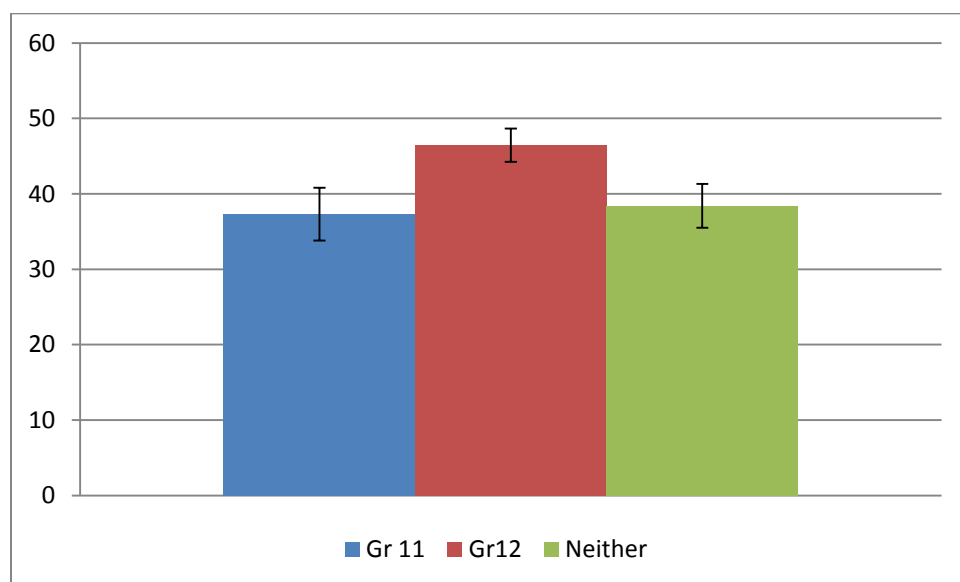


Figure 19: Post-course CLASS scores in both courses by level of high-school completed

4.6 Direct Correlation

In directly correlating the variables, very few trends were discovered. A moderate correlation was considered to be an absolute value between 0.3 and 0.7. It was expected that there would be a correlation between post-course CLASS scores and grades because this would indicate that the course evaluation methods were aligned with CLASS. There was a moderate correlation found in PHY100 (0.301), but not in PHY205 (0.244). It was also hypothesized that gain would correlate with grades, based on the assumption that more conscientious students would be more open to

adopting better attitudes towards learning physics. However, no correlation was found in either course. This indicates that students may succeed in these courses without necessarily meeting the intended goals of physics breadth courses.

Table 5: PHY100 correlation coefficients

Variable 1	Variable 2	Correlation Coefficient	Error
Post	Grades	0.301	0.125
Gain	Grades	-0.063	0.130

Table 6: PHY205 correlation coefficients

Variable 1	Variable 2	Correlation Coefficient	Error
Post	Grades	0.244	0.071
Gain	Grades	0.041	0.074

4.7 Specific Considerations in CLASS

Of particular interest were questions 8, 12, 16, 27, and 41 from CLASS, and one of the questions adapted from EBAPS. Questions 12 and 16 deal with students' outlooks on the limits of succeeding in physics. Ideally, students would disagree with Q12 and agree with Q16.

Q12. I cannot learn physics if the teacher does not explain things well in class.

Q16. Nearly everyone is capable of understanding physics if they work at it.

Students in both courses scored well on question 16 (41% and 45.7%), but not on question 12 (-47.5% and -31.2%) (Fig 20&21). These results suggest that students see success in physics as largely dependent on the lecturer. They think “nearly everyone is capable of understanding physics if they work at it . . .” but only if the instructor explains things well. Gray et al found that students did not succeed well in guessing the expert response to statement twelve.³ It would be interesting to see the responses of non-physics faculty because the issue of self-directed learning is not specific to physics. This statement might highlight a general difference between how students and professionals approach learning.

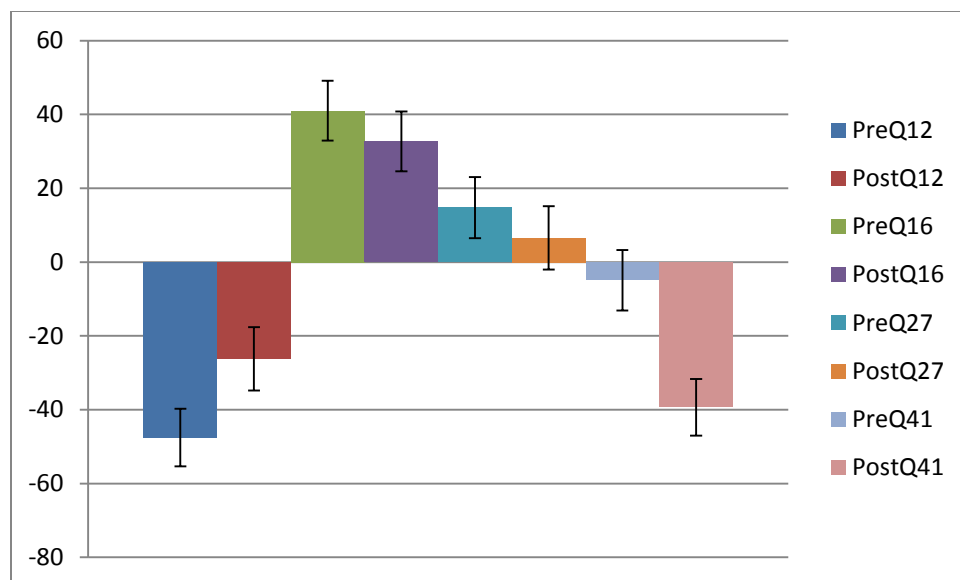


Figure 20: PHY100 pre- and post-course scores on selected questions

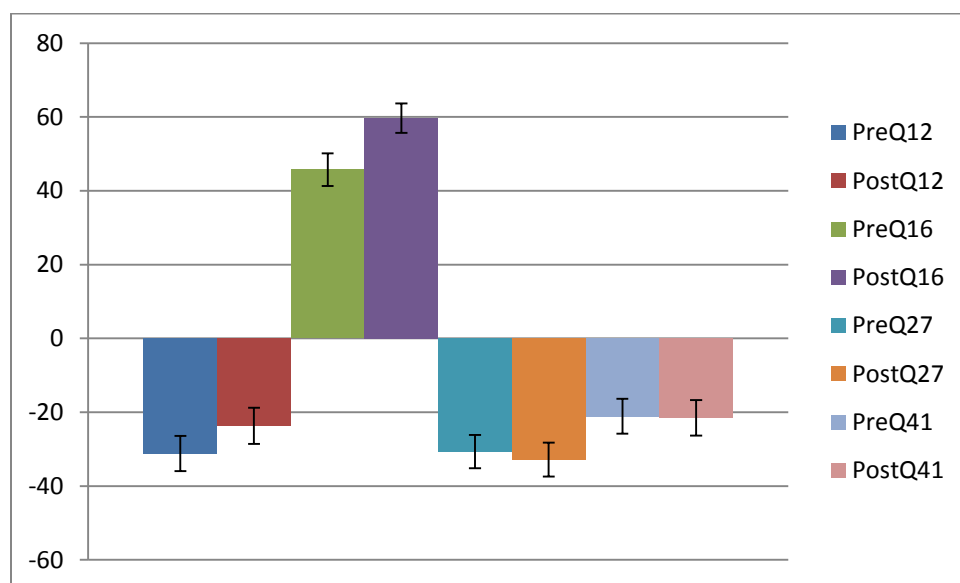


Figure 21: PHY205 pre- and post-course scores on selected questions

Questions 27 and 41, deal with the scientific process. Q27 is scored and Q41 is not scored, however, problems were found with both of them, stemming from the vague nature of the statements.

Q27. It is important for the government to approve new scientific ideas before they can be widely accepted.

Q41. It is possible for physicists to carefully perform the same experiment and get two very different results that are both correct.

Professionals unanimously agree on the interpretation of question 27 and disagree with the statement.³ In PHY100, marginally more students agreed with the statement than disagreed with it. In PHY205 the majority of responses to the statement were neutral (Fig 20&21). When students in the summer section of PHY205 were asked to respond in writing it became clear that students had no consensus on the interpretation of the statement. Many students interpreted the question as referring to how scientific theories should come to be widely, as opposed to how they do come to be widely accepted. Of these students, some of them disagreed with the statement (the expert response) because they think that politicians do not have the expertise to evaluate scientific work. There was also a sentiment of mistrust in the government, with some students referencing Galileo's legal troubles and the misuse of technology for political gain. The opposite sentiment was expressed by some students who agreed with the statement. They perceived scientists as amoral, and acting without regard to the societal impact of their research, which creates a need for the government to protect the public from discoveries run amok. Other students who agreed with the statement said that the government needs to help the public sort true science from pseudoscience. Some students took a practical interpretation of the statement and argued that it is easier for scientific ideas to be widely accepted if the government takes part in spreading scientific knowledge. Very few students mentioned the peer review process.

Looking at incoming students from different programs (Fig 22), it wasn't expected that there would be a large difference in responses to this question because very few undergraduate students have firsthand experience in research. However, arts and life science students showed a more expert-like view than did physical science or commerce students.

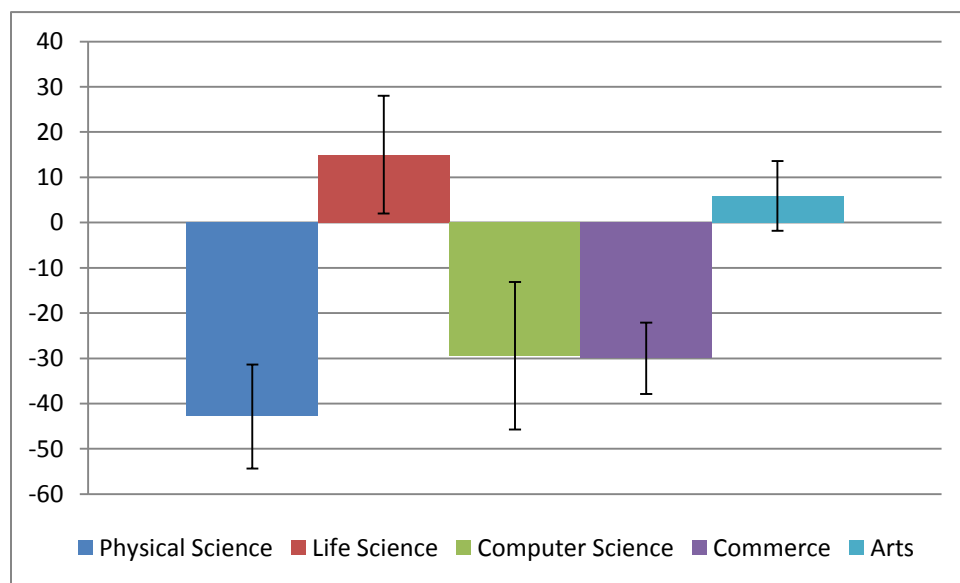


Figure 22: pre-course scores on question 27 in both courses by program of study

It was hypothesized that some of the ambiguity surrounding question 41 had to do with the interpretation of 'experimental results'. The students' limited understanding of quantum mechanics might confuse them about what constitutes a result. The results of experiments are typically the averages of many experimental measurements. This is especially true with experiments in quantum mechanics, where the underlying theory is statistical in nature. However, students may think of a result in science as the outcome of a single measurement. It is

possible that students with a rudimentary understanding of non-locality think of the double slit experiment as an example of two different, but correct, results. They are taking *result* to mean the location of one electron on the screen, as opposed to the distribution of many electron locations.

This hypothesis was supported by the data from this study (Fig 23). The students in PHY205 only learned classical physics and their responses to Q41 did not change significantly. The students in PHY100 did learn basic concepts of quantum mechanics, including the uncertainty principle, non-locality, and the double slit experiment. Their average score on Q41 became significantly more novice-like over the term.

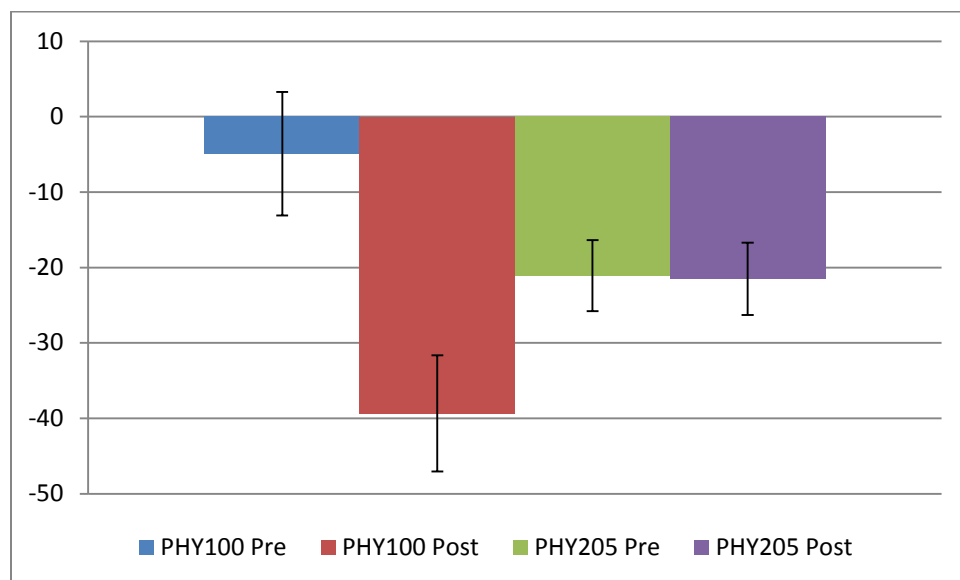


Figure 23: pre- and post-course scores on question 41 in each course

Three questions were added from EBAPS on the pre-course survey, and several questions were added on the post-course survey, asking about students' experience in the course:

EBAPS:

1. Computer simulations can predict the behavior of physical objects like comets. But simulations can also help scientists estimate things involving the behavior of *people*, such as how many people will buy new television sets next year.⁷
2. Understanding science is really important for people who design rockets, but not important for politicians.⁷
3. When it comes to controversial topics such as which foods cause cancer, there's no way for scientists to evaluate which scientific studies are the best.⁷

Post-Course:

43. This class has been a positive experience for me.
44. Students in my program should be required to take a science class.

The expert answer is to agree with question one, and disagree with two and three.⁷ Questions 43 and 44 don't have expert opinions, but hopefully students agree with both of them. Questions 2 and 44 are similar. Literal interpretations aside, question 2 asks if students think that non-scientists should have an understanding of science and question 44 asks if students think they themselves should have an understanding of science. Students scored more poorly on 44 than on two (Fig 24&25), indicating that some think scientific literacy is generally important, except when it comes to themselves.

Questions one and three both address the nature of science as a method of inquiry, and were answered fairly well in both courses. Question 43 was answered positively, which shows that both courses tend to give students a good association with science. This is an important accomplishment given the mistrust of scientists that some students showed in their responses to question 27.

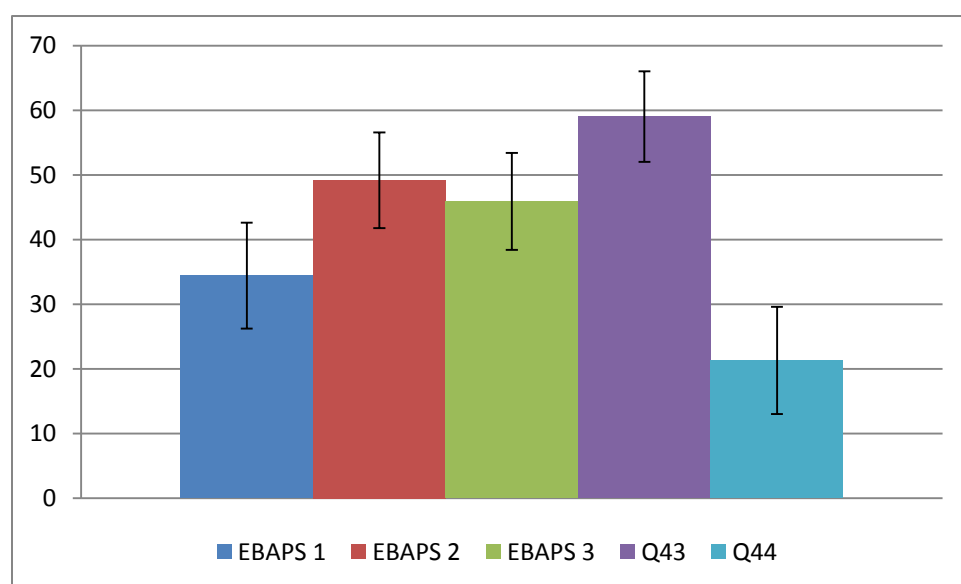


Figure 24: PHY100 scores on additional questions

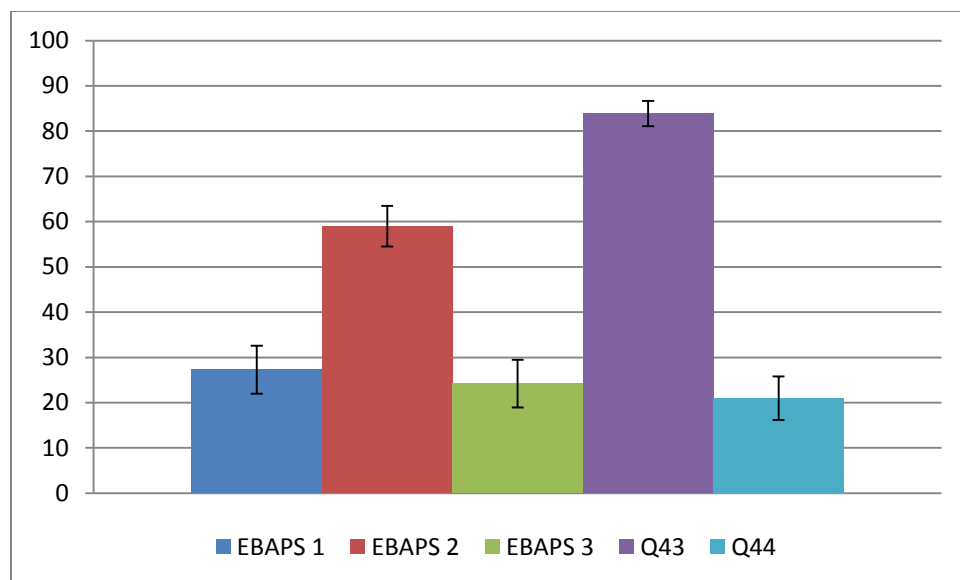


Figure 25: PHY205 scores on additional questions

Most students do not do well on question eight of CLASS, which reads: “when I solve a physics problem, I locate an equation that uses the variables given in the problem and plug in the values.” This statement describes the ‘plug and chug’ approach to problem solving, which relies on algebraic manipulation as opposed to understanding concepts. Although solving physics problems often includes the above approach, experts will begin solving a problem by picturing what is physically happening before making use of equations. Gray et al found that students did not do well at guessing the expert response to this statement.³ This is not surprising because this method of problem solving is taught in some high-school math classes.

When comparing incoming students (Fig 26), computer science students scored the highest (although still negative) and life science students scored the lowest. When students were asked to solve interview question one, arts students used the ‘plug and chug’ method, while the computer science student drew a free body diagram, stated the assumptions made, and talked through the calculation. When asked about his/her program, (s)he revealed that many computer science students have extra-curricular computational projects, and this constantly challenges them to be innovative and test their understanding of material learned in class.

Eighty percent of life science students agreed with the statement and none of them disagreed with it. Many of them are intending to pursue medicine which requires very competitive GPA’s. This simplistic method of problem solving may be especially ingrained in students whose end goal is an easy ‘A’ and not a thorough understanding of material. This inference is supported by the findings of Adams et al.² Students who were interviewed about sense-making and effort labeled themselves as ‘lazy’ in their problem solving approaches. They expect physicists to think carefully about physics problems, but they themselves want to optimize their study time while balancing five courses. This means that if students can succeed in a course without working to make sense of the material then some of them will.

With regards to question eight, students in PHY100 and PHY205 who were expert-like did no better than novice-like students (Fig 27). Both courses directly evaluate conceptual

understanding. It is possible that students do put effort into understanding but do not apply it to problem solving.

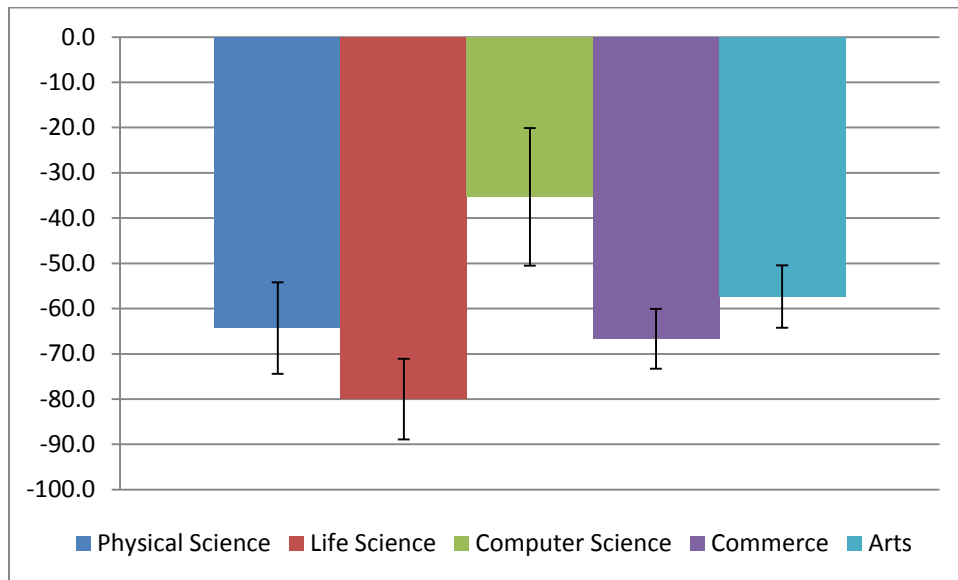


Figure 26: Pre-course scores on question 8 in both courses by program of study

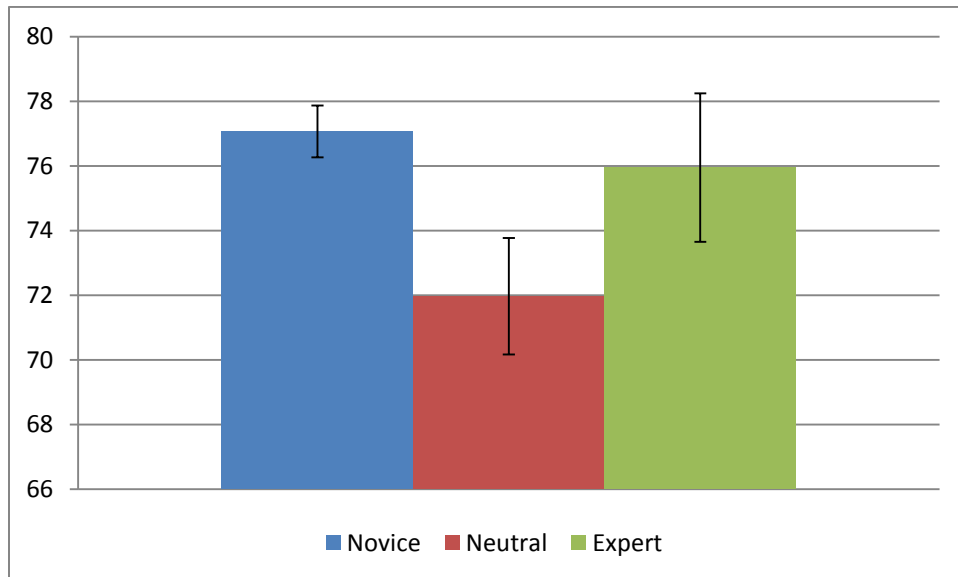


Figure 27: Final course grade in both courses by response to question 8

5. Conclusion and Future Work

Neither course significantly altered students' attitudes towards learning physics overall. Almost as many student CLASS scores decreased over the term as increased. Furthermore, the students with higher final scores did not necessarily obtain better grades in the course, as there was only a moderate correlation between post-course score and grades. As at other institutions, the best scored categories were *real world connections*, *personal interest*, and *sense making/effort*. The worst scored were *applied conceptual understanding* and *problem solving sophistication*. This suggests that students who choose to satisfy breadth requirements with physics courses are truly interested in the material instead of perceiving the subject as an easy 'A'. It seems that students can learn algorithms to solve algebraic physics problems, and can also learn isolated physical concepts, but have trouble combining various aspects of learning physics.

PHY100 appears to provide a more even playing field for students from various backgrounds. Neither program of study nor high-school experience significantly affected grades, post-course CLASS scores, or CLASS gain, despite incoming arts students scoring lower on pre-course surveys. PHY100 requires students to write, which is a skill that arts students claimed to possess and use in their own programs. The writing requirements in the course might contribute to the success of arts students, but there is a possibility that it disadvantages ESL students. This could be tested in further work by including a question about first language on the surveys.

The success of students in PHY205 was somewhat dependent on program of study and high-school experience. Science students made up 38% of the course population, which might have created problems for the arts students. Arts students had the lowest grades and post-course scores, both significantly worse than commerce students. The increasingly quantitative nature of this course might be due to feedback from students who are not taking it as a breadth requirement. PHY205 students significantly increased their scores in *problem solving confidence* and *problem solving sophistication*, which appears to be a positive effect of the emphasis on calculations.

To save time in future survey analysis, the original list of student responses should only include students who wrote both surveys, correctly answered question 31, and properly indicated their program of study and previous physics education. This initial step will avoid unnecessary complications and possible mistakes later on.

Out of the 516 surveys submitted only 245 were used. This huge loss can be attributed to the requirement for submission of both pre- and post- course surveys, and the omission of surveys from students who chose 'Other' as their major. Loss could have been avoided by not giving students the option of 'Other'. Additionally, when indicating their programs, some students answered both parts A and B, despite being told to only answer one of them. It would have been less confusing to contain all of the options in one question. This can be done using the five groupings used in these results, with the addition of 'fine arts' as its own category. We did not have enough fine arts students to treat them in isolation, but our preliminary data suggests they perform distinctly. It is also advisable to address the issue of double majors so as to avoid confusion.

The creators of CLASS claim that it should take students 10 minutes or fewer to fill out the survey², although it took 15-20 minutes for students at UofT to complete the survey. If this is too

much time to take out of tutorials, students could be offered an incentive to complete the surveys online. Interviewed students said CLASS was somewhat long and repetitive. A possibility is to administer CLASS with only the 36 score-able questions until the next version is developed. Out of the added questions from EBAPS, only the second one was revealing.

The research summarized in the introduction suggests that attitudes can be improved through either explicit discussion or implicit directed learning. If instructors wish to improve scores in certain areas of CLASS, they can explicitly discuss the issues addressed by CLASS statements they find important. Some students in PHY205 thought that the tutorial experiments were too predictable. If these experiments were designed so that students felt like they were discovering new concepts for themselves, the tutorials might act similarly to the PbI curriculum, which has been shown to improve student performance on CLASS and the FCI. Many CLASS statements address the goals laid out for breadth courses in the curriculum document, including the exposure to physics methodology and fostering of interdisciplinary thinking. Therefore, CLASS is a relevant measure of the effect of physics breadth courses.

Acknowledgments

Thank you to April Seeley for processing the surveys, and to all of the students who participated in the surveys and interviews.

Sources:

1. University of Toronto, Faculty of Arts and Science, Degree Objectives Committee., Proposal for New Breadth Requirement, (2009)
2. W. K. Adams et al., New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. 2, 010101 (2006)
3. K. E. Gray et al., Students know what physicists believe, but they don't agree: A study using the CLASS survey, Phys. Rev. ST Phys. Educ. Res. 4, 020106 (2008)
4. Valerie K. Otero and Kara E. Gray., Attitudinal gains across multiple universities using the Physics and Everyday Thinking Curriculum, Phys. Rev. ST Phys. Educ. Res. 4, 020104 (2008)
5. B. A. Lindsey et al., Positive attitudinal shifts with the Physics by Inquiry curriculum across multiple implementations, Phys. Rev. ST Phys. Educ. Res. 8, 010102 (2012)
6. M. Milner-Bolotin et al., Attitudes about science and conceptual physics learning in university introductory physics courses, Phys. Rev. ST Phys. Educ. Res. 7, 020107 (2011)
7. http://www2.physics.umd.edu/~elby/EBAPS/EBAPS_items.htm

Appendix A: CLASS Questions²

1. A significant problem in learning physics is being able to memorize all the information I need to know.
2. When I am solving a physics problem, I try to decide what would be a reasonable value for the answer.
3. I think about the physics I experience in everyday life.
4. It is useful for me to do lots and lots of problems when learning physics.
5. After I study a topic in physics and feel that I understand it, I have difficulty solving problems on the same topic.
6. Knowledge in physics consists of many disconnected topics.
7. As physicists learn more, most physics ideas we use today are likely to be proven wrong.
8. When I solve a physics problem, I locate an equation that uses the variables given in the problem and plug in the values.
9. I find that reading the text in detail is a good way for me to learn physics.
10. There is usually only one correct approach to solving a physics problem.
11. I am not satisfied until I understand why something works the way it does.
12. I cannot learn physics if the teacher does not explain things well in class.
13. I do not expect physics equations to help my understanding of the ideas; they are just for doing calculations.
14. I study physics to learn knowledge that will be useful in my life outside of school.
15. If I get stuck on a physics problem on my first try, I usually try to figure out a different way that works.
16. Nearly everyone is capable of understanding physics if they work at it.
17. Understanding physics basically means being able to recall something you've read or been shown.
18. There could be two different correct values to a physics problem if I use two different approaches.
19. To understand physics I discuss it with friends and other students.
20. I do not spend more than five minutes stuck on a physics problem before giving up or seeking help from someone else.
21. If I don't remember a particular equation needed to solve a problem on an exam, there's nothing much I can do (legally!) to come up with it.
22. If I want to apply a method used for solving one physics problem to another problem, the problems must involve very similar situations.
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.
24. In physics, it is important for me to make sense out of formulas before I can use them correctly.
25. I enjoy solving physics problems.
26. In physics, mathematical formulas express meaningful relationships among measurable quantities.
27. It is important for the government to approve new scientific ideas before they can be widely accepted.
28. Learning physics changes my ideas about how the world works.
29. To learn physics, I only need to memorize solutions to sample problems.
30. Reasoning skills used to understand physics can be helpful to me in my everyday life.

31. We use this question to discard the survey of people who are not reading the statements. Please select agree - option 4 (not strongly agree) to preserve your answers.
32. Spending a lot of time understanding where formulas come from is a waste of time.
33. I find carefully analyzing only a few problems in detail is a good way for me to learn physics.
34. I can usually figure out a way to solve physics problems.
35. The subject of physics has little relation to what I experience in the real world.
36. There are times I solve a physics problem more than one way to help my understanding.
37. To understand physics, I sometimes think about my personal experiences and relate them to the topic being analyzed.
38. It is possible to explain physics ideas without mathematical formulas.
39. When I solve a physics problem, I explicitly think about which physics ideas apply to the problem.
40. If I get stuck on a physics problem, there is no chance I'll figure it out on my own.
41. It is possible for physicists to carefully perform the same experiment and get two very different results that are both correct.
42. When studying physics, I relate the important information to what I already know rather than just memorizing it the way it is presented.

Appendix B: CLASS Categories ¹

Real World Connection 28, 30, 35, 37
 Personal Interest 3, 11, 14, 25, 28, 30
 Sense Making/Effort 11, 23, 24, 32, 36, 39, 42
 Conceptual Connections 1, 5, 6, 13, 21, 32
 Applied Conceptual Understanding 1, 5, 6, 8, 21, 22, 40
 Problem Solving General 13, 15, 16, 25, 26, 34, 40, 42
 Problem Solving Confidence 15, 16, 34, 40
 Problem Solving Sophistication 5, 21, 22, 25, 34, 40
 Not Scored 4, 7, 9, 31, 33, 41

Appendix C: Interview Problem Solving Questions

1. A penny of 0.1kg is dropped from a height of $d = 2000\text{m}$. Initially, it has speed $v = 0\text{ m/s}$ and total energy $E = 2000\text{J}$. Its diameter is 2cm. Take the acceleration due to gravity to be $g = 10\text{m/s}^2$. Neglect air resistance.
 $d = \text{acceleration} \cdot \text{time}^2 / 2$
 $E_G = \text{height} \cdot \text{mass} \cdot g$
 $v_{\text{final}} = \text{acceleration} \cdot \text{time}$
 $E_{\text{kinetic}} = m \cdot v^2 / 2$
 What is the final speed of the penny right before it hits the ground?
2. A kid tries to paddle a canoe by moving the paddle backwards and forwards in the water. The canoe moves a couple of inches back and forth with the paddling, but does not travel forwards significantly. The kid's parent suggests taking the paddle out of the water after every stroke to propel the canoe forwards. Why do you think this will work?
3. If $F_{\text{electric}} = k \cdot q_1 \cdot q_2 / d^2$, what are the units of k ?
4. Why do sharks and airplanes look similar?