Final report on CA3 assessment academic year 2009/2010

Prepared by

Annelie Skoog

Associate Professor

Department of Marine Sciences

August 2010

1 BACKGROUND

During academic year 07/08, Hedley Freake and Elisabeth Kloeblen carried out a survey of coverage of CA3 learning goals in 9 courses and Scott Brown also put together a student self-efficacy instrument. With slight modifications by Scott Brown and Annelie Skoog in 2009, the student self-efficacy instrument was used pre- and post-course in spring 2009 as an on-line quiz. A very large data set was collected, which was only partially analyzed during spring/early summer 2009.

The conclusions and recommendations from the spring 2009 assessment report included recommended changes to the self-efficacy instrument, a more in-depth study of negative student responses to laboratory sections, implementation of a student-learning survey during spring 2010 and reporting of the results from the survey during a workshop. Two main topics were addressed in the spring 2010 assessment: further evaluation of the large data set on student self-efficacy from spring 2009 and implementation of a survey on actual student-learning.

2 ASSESSMENT OF ACTUAL STUDENT LEARNING

2.1 Methods

2.1.1 Assessment participation

The focus group for the request to participate in the 2010 survey was the same group of faculty participating in the self-efficacy assessment in 2009. The survey on actual student-learning was met with enthusiasm when it was first brought up in spring 2009. However, the actual participation in the survey in 2010 was very low, and it is not quite clear why. The courses participating in the assessment were: GEOG2300, MARN1002, MARN1003, and PSYCH1100, with enrollments of 108, 81, 37 and 276, respectively.

The survey was proposed to be part of the final exam as questions written by each faculty member and geared towards their specific topic. It is possible that the faculty found the actual writing of these questions to be an obstacle. A potential reason may be that the learning outcomes are somewhat unclear. This potential problem will be assessed at a workshop fall 2010.

2.1.2 Assessment methods

2.1.2.1 GEOG2300

The faculty member teaching GEOG2300 assessed learning goals 1, 2, and 3 (Table 1) as part of a multiple-choice, comprehensive final. In this course, the final is optional - students wishing to improve their course grade can take the final, no one else has to. When asked how many students actually took the final, the faculty member was unsure. When calculating averages, the highest number of correct answers on a question (53) was used as the number of students taking the final.

2.1.2.2 MARN1002 and 1003

The faculty member teaching MARN1002 and 1003 divided learning goal 1 in two parts: 1a) Know basic concepts and vocabulary of oceanography; 1b) Know the importance of oceanography for society. These two learning goals were assessed separately. All learning goals were assessed as part of a comprehensive final using multiple-choice questions. Averages were calculated by dividing the number of correct answers by the total number of students taking the exam.

2.1.2.3 PSYCH1100

The faculty member teaching PSYCH1100 provided a lot of data. He went through all three tests given during the semester (all multiple choice) and assigned multiple learning goals to each question. The number of students taking exams 1, 2, and 3 were 276, 263, and 224, respectively. These exams were all given as bubble-sheet, multiple-choice exams and the exams were corrected centrally. When examining the data closely, it became clear that the system calculates average result for a question by dividing the number of correct answers with the number of students listing an answer for the question. Note that this is different from the calculation for averages in GEOG2300, MARN1002, and MARN1003 where averages were calculated by dividing the number of correct answers by the number of students taking the exam.

Table 1. Learning goals for content area 3 courses.

Number	Learning goal
1	Know the basic concepts and vocabulary of two areas of science or technology and the importance of these areas to modern society
2	Be familiar with at least two contemporary scientific or technical methods and understand how they are applied to gain scientific or technical knowledge.
3	Be able to explain the conceptual basis of the Scientific Method , including its definition, motivation, steps of application, hypothesis testing, and misapplications.
4	Be able to distinguish between science and pseudoscience.
5	Be able to describe a scientific experiment that he or she is familiar with and explain how it applies the steps of the scientific method.
6	Be familiar with some unresolved scientific questions.
7	Be able to analyze debates about the roles science and technology play in shaping the world and human society.
8	Acquire skills associated with scientific inquiry.

2.2 Results and discussion

Interestingly, and perhaps unexpectedly, learning outcome 1 had a relatively low average. This learning outcome incorporates two rather different parts - concepts and vocabulary of a specific field and the importance of this field to society. It seems this outcome should be relatively easy for students, since it involves route learning, which has been shown to be much easier for students than critical thinking. Critical thinking would be incorporated in learning outcomes 4 and 7, which both had higher results than learning outcome 1. For MARN1002 and 1003, learning outcome 1 was divided in two parts: 1a) Know basic concepts and vocabulary of oceanography; 1b) Know the importance of oceanography for society. The result for MARN1002 was 73% and 98%, respectively, and the result for MARN1003 was 70% and 95% respectively. The averages (85% and 82%) are shown in Fig. 1.

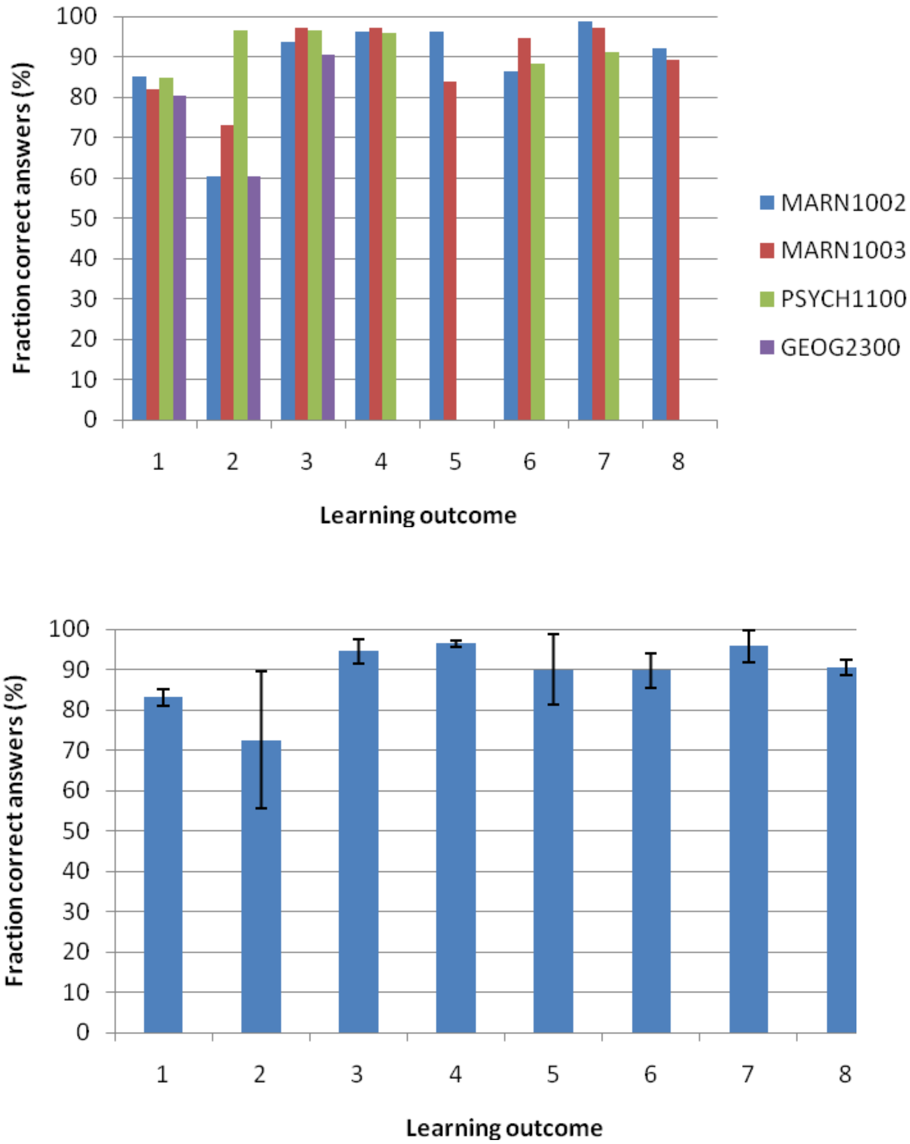

Figure 1. A. Results from final exams for GEOG2300, MARN1002, and MARN1003. For PSYCH1100, the results are from the last exam of the semester. B. Average fraction of correct answers calculated from course results in Fig. 1A. Error bars denote one standard deviation.

Learning outcome 2 had the lowest average and the largest difference between courses. The technical understanding seems lower in the MARN and GEOG courses than in psychology. It is difficult to speculate about a reason, but perhaps the more technological characteristics of the oceanography and geography field make it harder for students to grasp methods associated with these fields.

Another unexpected result is the high scores associated with learning outcomes pertaining to the scientific method (outcome 3, 4, and 5). In the assessment report for AY 07/08, Freake and Kloeblen found that few courses actively taught the scientific method. The data presented for 09/10 indicates that students have a good grasp of the scientific method. However, since the sample size for 09/10 is very small it can't be said for certain whether students in general actually have this knowledge. From asking students in my courses (MARN1002 and 1003) I also found that most students have had exposure to the scientific method in high school and middle school. This may support the notion that most students know of and understand the scientific method by the time they have finished their CA3 course requirement. However, it may or may not be the result of what they have learned at UConn.

Data from PSYCH1100 show a very positive trend (Fig. 2) - there is a steady increase in average results for most learning outcomes as the semester progresses. However, note that the number of students in PSYCH1100 decreased from 276 to 225 from exam 1 to exam 3. Weaker students dropping the course could account for some of the apparent improvement.

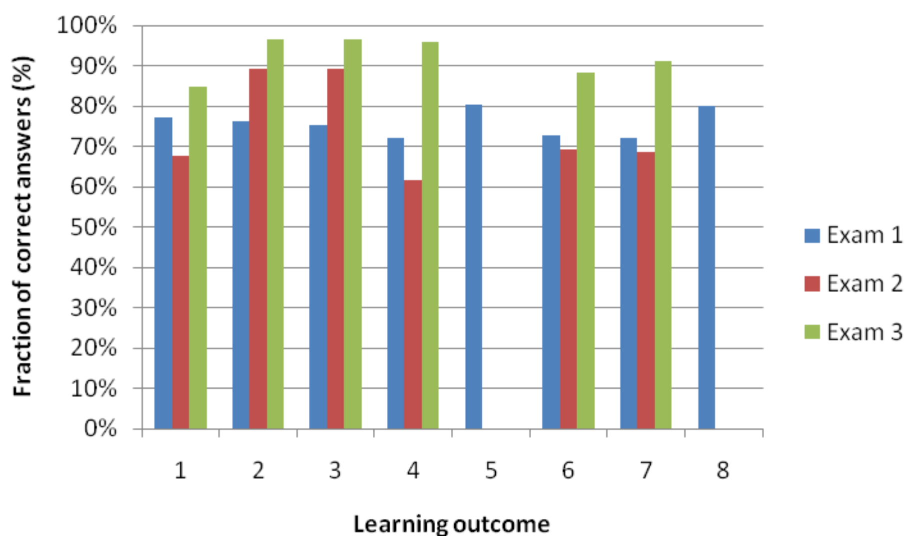

Figure 2. Data from PSYCH1100. n is 276, 263, and 225 for exams 1, 2, and 3, respectively.

2.2.1 Interpretation of learning goals

When attempting to assess the learning goals as part of my final in MARN1002 and 1003, I found some learning goals difficult to interpret. This is a critical flaw, since one of the most important characteristics of learning outcomes is that they should be clear and measurable. I found some learning goals unclear, some unrelated goals are paired as one outcome, and some learning goals are redundant. Learning goal 1 pairs two unrelated learning outcomes - knowing basic concepts and vocabulary of a field is paired with knowing the importance of a specific field to society. It is possible that the main point of the learning outcome is to state that knowledge of two scientific fields is required, i.e. in order to fulfill the general education requirements a student has to take two science courses. If so, this could be expressed more clearly. Learning goal 4 should be part of learning goal 3 - it makes no sense to have a separate learning goal on pseudoscience since the definition of pseudoscience is "knowledge" not based on the scientific method. Learning goal 5 is simply a misrepresentation of the scientific method - the scientific method is not applied to one experiment but to a series of experiments. This learning goal also seems redundant and should be incorporated with learning goal 3. Learning goal 7 is very vague. What does analyze debates refer to? If "analyzing debates" refers to differentiating between parts of a debate based on science and parts based on pseudoscience, that has already been covered. I interpreted learning outcome 8 to mean that students should be able to use methodology used in oceanography. This could refer to the lab course, but it could also refer to basic skills in math and understanding graphs.

It may not be possible to change the CA3 learning outcomes at this stage, but they should be discussed as part of the workshop planned for the fall. For a future assessment effort it would be useful to clarify what each statement means - this would make it easier for faculty to incorporate assessment questions in their exams. I'm assuming there was a committee responsible for assembling these outcomes and it would be interesting to have one or more committee members be part of the discussion.

2.2.2 How to calculate an average?

How to calculate an average may seem obvious, but in evaluating the student learning assessment I learned there are two possibilities: 1) The average is calculated by dividing the total number of correct answers with the total number of students taking the exam or 2) The average is calculated by dividing the total number of correct answers with the total number of students answering the question.

The data given by the system that corrects bubble sheet exams is based on an average calculated from the number of students answering the question. This would always give a higher than or equal average when compared to an average calculated from the total number of students taking the exam. In the context of assessment, it can be argued that by choosing not to answer the question, the students have shown that they don't know the answer. In which case the average should reflect that the students don't know. On the other hand, it also frequently happens that students run out of time on an exam, in which case an unanswered question should not be interpreted as the student not knowing the answer. What to do? This question should probably come up during the seminar on assessment in the fall.

3 FURTHER EVALUATION OF PRE- AND POST-COURSE STUDENT SCIENCE SELF-EFFICACY STUDY CARRIED OUT AY 2008/2009

3.1 Data gathering and evaluation methods

Detailed description of methods can be found in "Final report on CA3 assessment AY 2008/2009" by Annelie Skoog. Briefly, 32 courses participated in the on-line survey, which was delivered pre- and post-course. The survey contained 13 statements (Table 2) rated on a 5-step scale from strongly disagree (given a numerical value of 1) to strongly agree (given a numerical value of 5).

During the assessment effort 2008/2009, all collected data were used. In contrast, during the further evaluation in 2009/2010, only data points where the same student had supplied both pre- and post-course answers were used (denoted paired data in the report from 2008/2009). Since a paired t-test needs the same number of data points in both data arrays used in the test, if any answer was missing from either assessment, this student's data was excluded from the evaluation. Employing this criterion resulted in 613 students used in both the pre- and post-course data arrays. During the assessment effort

2008/2009, there was no difference between averages of data from sample sizes of ~1200 students versus ~700, indicating that a sample size <= 700 students is sufficient.

A student's pre-course rating of a statement was subtracted from the post-course rating. If the resulting value is positive it can be interpreted as an improvement in self efficacy. The average for each statement was then calculated. Note that question 11 (see Table 2) is different - an increase in science self efficacy would result in a negative value.

Unless otherwise noted, statistically significant refers to a probability < 0.05.

Table 2. Statements used in the pre and post-course assessments spring 2009. Statements 11 and 13 were the same in the pre- and post-course assessment.

Pre-course: I am confident that I can answer questions on: Post-course: After taking a CA3 Course, I am confident that I can answer questions on:
1. Basic concepts and vocabulary taught in the course
2. The methods and technologies utilized by scientists in the discipline
3. The application of the Scientific Method
4. The difference between science and pseudoscience
5. The conduct of a scientific experiment I am familiar with
6. The identity of unresolved questions in the field of science
7. How science impacts society
8. Pre-course: I am confident that I can apply my science knowledge to events in everyday life Post-course: After taking a CA3 Course, I am confident that I can apply my science knowledge to events in everyday life
9. Pre-course: By taking a lab course, I will improve my practical science skills Post-course: By taking a lab course, I improved my practical science skills
10. Pre-course: I like science Post-course: I like science more after taking a CA3 Course
11. I find it difficult to understand current scientific events in the news
12. Pre-course: I am interested in science Post-course: After taking this CA3 Course, I am more interested in science
13. I will likely seek out more information about science through (check all that apply) _ Another course_ Internet_ News/ Media_Other: __ I will not seek out more information

Tested hypotheses include:

- Low-enrollment courses give larger improvements in average science self-efficacy than high-enrollment courses.
- Female students in low-enrollment courses have larger improvement in average self-efficacy results than male students in low-enrollment courses.
- Students with high GPA have higher average self-efficacy results than students with low GPA.
- Students with low GPA have a larger improvement in average self-efficacy results than students with high GPA.

3.2 Results and discussion

3.2.1 General trends in paired data

Last year, we found that students were fairly confident about their science skills even before taking courses, evidenced by the large number of responses with averages close to 4, that is, a verbal response of "agree". Paired data (Fig. 3) gave statistically significant improvements (paired t-test for means) only for statement 2, 4, and 6, and the improvements were very small. Actual numbers for the improvements were 0.13, 0.32, and 0.11, corresponding to 2.5%, 6.4%, 2.1%, respectively.

Disappointingly, students did not find that taking a laboratory course improved their practical science skills (statement 9, Fig. 1), nor had the courses improved like of (statement 10) or interest in (statement 12) science. Similarly, when using pooled data last year, the same statements had significant differences (two-tailed t-test, equal variances), and the direction of change was also the same.

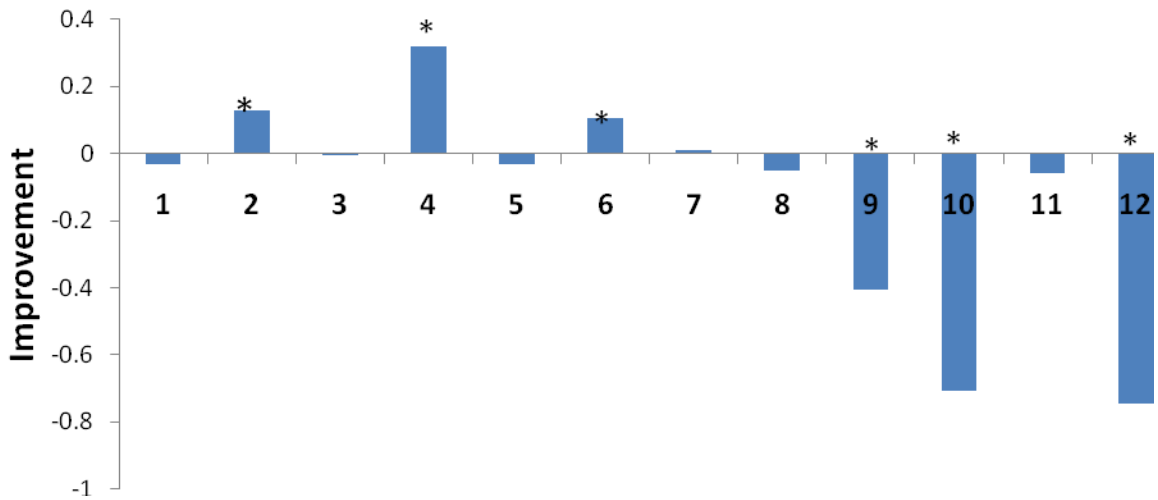

Figure 3. Averages in differences between post- and pre-course statements using all paired data. With exception of statement 11, a positive value would indicate a perceived improvement. Labels on x-axis refer to statements 1-12 in table 1. Asterisks denote statistically significant differences between pre and post-course results. N is 613 for both the pre- and post-course assessment.

3.2.2 Gender differences

Gender differences were analyzed using paired data from 283 females and 329 males. When using paired data, statement 1 had the only statistically significant difference between female and male students (data not shown) - male students felt they had improved their knowledge of concepts and vocabulary by an average of 1% by taking the course, while, on average, female students did not perceive any improvement.

When using all data last year, male students (n= 702) were slightly more confident (two-tailed t-test) than female students (n= 668) in their pre-course science ability in 6 out of 12 areas (2, 4, 6, 7, 10, and 11). The results from the paired data reported above indicate there was none or a very small improvement by taking the course. In contrast, the post-course analysis of pooled data showed that the difference between males and females decreased; post-course, male students (n=520) were more confident than female students (n= 494) in only 3 (4, 6, and 9) out of 12 areas (two-tailed t-test). Therefore, it appears that the no-change implied by the paired data does not agree with the analysis of the pooled post-course data. However, the differences calculated from the pooled data are small, and it is likely that the smaller number of paired data had insufficient statistical power to detect this small change.

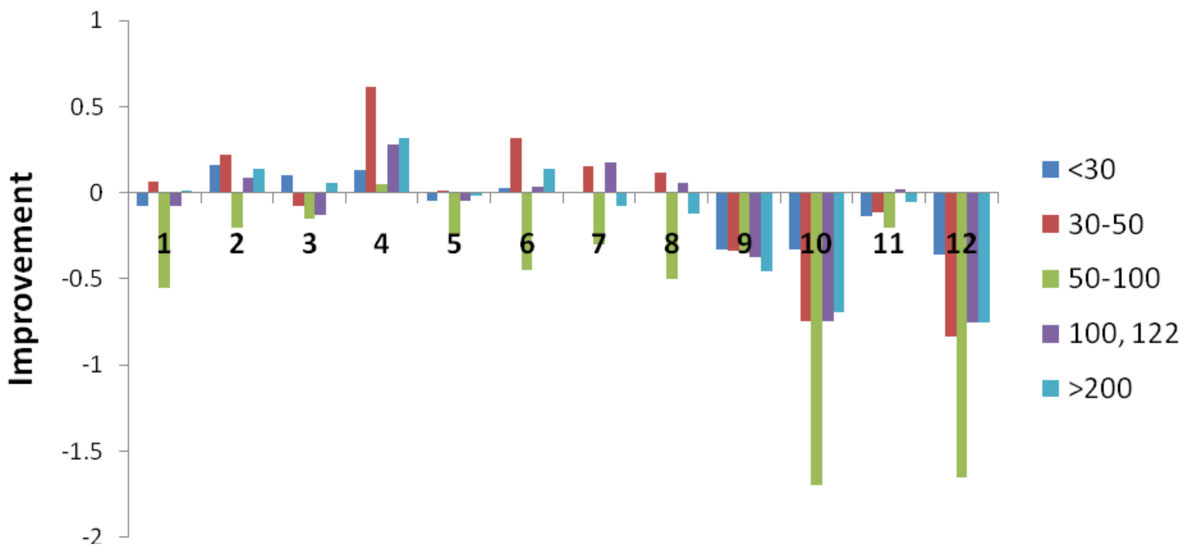

Figure 4. Effect of class size on perceived improvement in science knowledge. Labels on x-axis refer to statements 1-12 in table 1. The 50-100 students class size is significantly different from all other class sizes for statements 1, 2, 6, 7, 8, 10, and 12.

3.2.3 Effect of class size

There was no clear overall pattern in the effect of class size (Fig. 4). Courses with 50-100 students were significantly different from all other class sizes for statements 1, 2, 6, 7, 8, 10, and 12. However, this subgroup consisted of only two courses, and the difference may be a result of characteristics other than class size.

3.2.4 Connection between science self efficacy and grade point average

Interesting trends emerged when data was evaluated in bins based on GPA (Fig. 5). GPA including spring 2009 for individual students was retrieved from Peoplesoft. Before taking a CA3 course, like of (statement 10) and interest in (statement 12) science was significantly correlated ($p<0.05$) with GPA. Correlation coefficients were 0.91 and 0.89, respectively, indicating that GPA can explain 82% and 79% of the variation in how much students liked and were interested in science. Pre-course, none of the other statements had significant correlations with GPA.

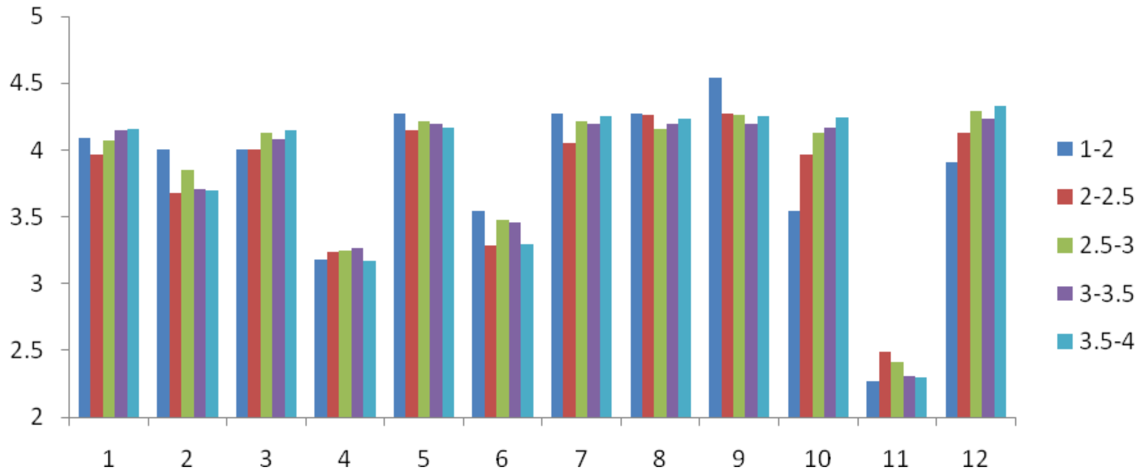

Figure 5. Pre-course average values for statements 1-12 sorted by GPA. Correlations between GPA and values for statements 10 and 12 are statistically significant.

After the course, statements 10 and 12 were still significantly correlated with GPA with correlation coefficients similar to the pre-course data. In addition, statements 3, 4, and 5 were also significantly positively correlated with GPA after the course (Fig. 6). It is interesting to note that student evaluations of science self efficacy were less correlated with GPA before the course than after the course, and an interpretation of that change is not self evident.

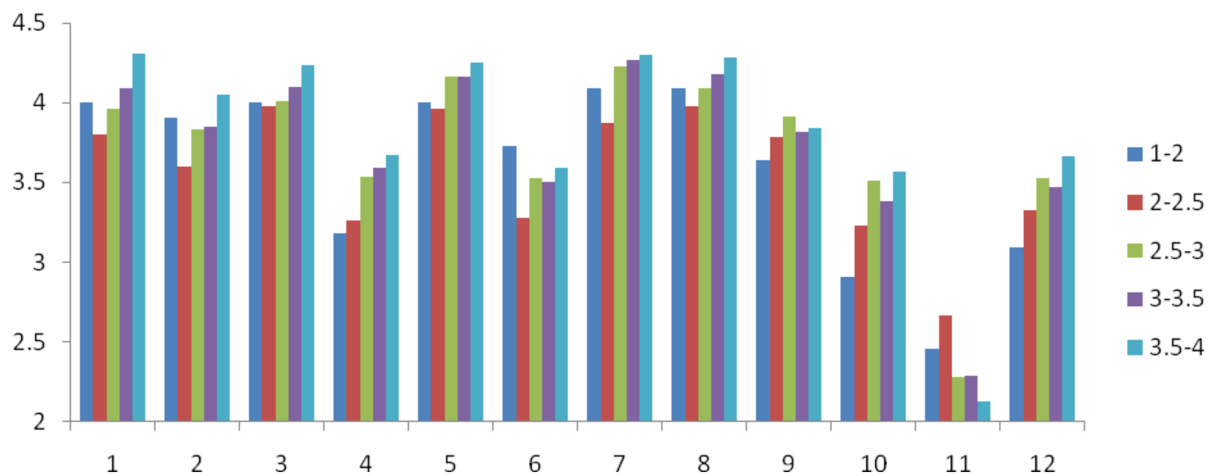

Figure 6. Average values for statements 1-12 post-course sorted by GPA. Correlations between GPA and values for statements 3, 4, 5, 10, and 12 are statistically significant.

Disappointingly, students with high GPA had higher average self-efficacy increases than students with low GPA for statements 2, 4, 5, 7, and 12 (Fig. 7). Apparently, we are not effectively reaching and teaching science and technology to students with a low GPA.

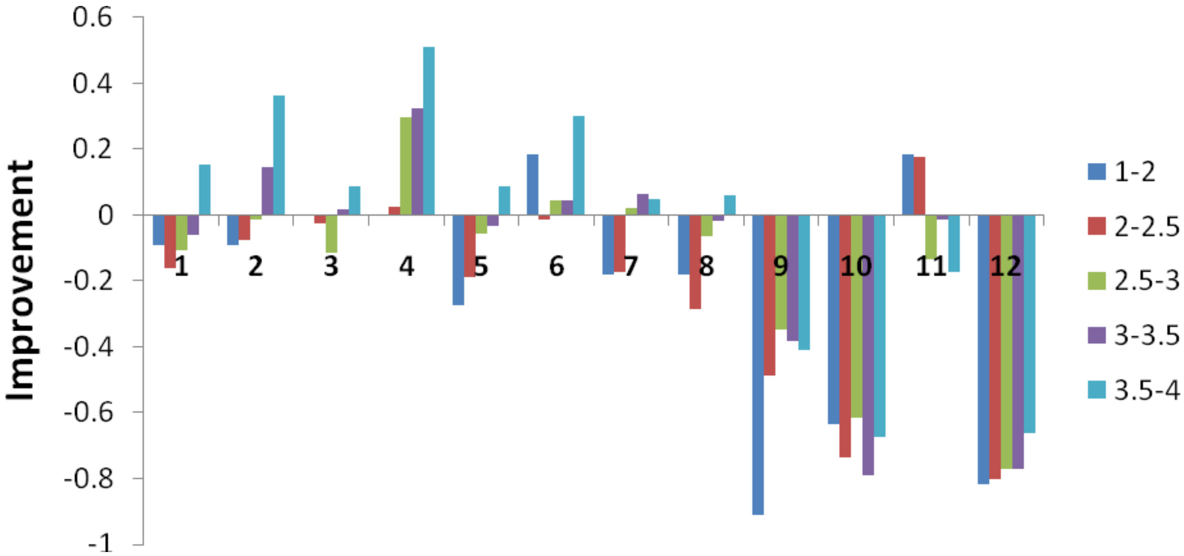

Figure 7. Improvement in understanding after taking a CA3 course, binned by GPA.

4 CONCLUSIONS

4.1 Assessing actual student learning

- Learning goals were somewhat difficult to interpret. In order to make it easier for faculty to participate in a future assessment of actual student learning, it would be necessary to make the learning goals easier to understand. This could be part of the workshop planned for fall 2010.
- Participants in this year's assessment of student learning were few, but despite the low participation rate we learned that assigning learning goals to questions can be done two ways - each question can assess only one learning goal or multiple learning goals. This could also be a methodological question to bring up at the fall work shop.
- Results were generally high, with average values of 73% to 96% calculated from the four participating courses.
- Student learning showed the lowest results for learning goal 2, implying that that understanding technologies and their applications is difficult for students.
- Students appear to have a good grasp of the scientific method and its applications, which is contrary to the fact that few CA3 courses actively teach the scientific method. However, note the low participation in this year's assessment.
- PSYCH1100 evaluated a series of three exams, which encouragingly showed increasing understanding of several learning outcomes as the semester progressed.

4.2 Student self efficacy

- There was no pattern indicating that class size affects improvements in self efficacy based on paired data.
- There were no statistically significant differences between males and females in improvements in self efficacy based on paired data. Note that pooled data indicated that males were more confident in their science abilities than females before taking a CA3 course. Further, this gender difference in self-efficacy decreased after taking a CA3 course based on pooled data.
- Students with a high GPA agreed that they liked and were interested in science (value close to 4) before taking a CA3 course, while students with a lower GPA were more neutral (average value slightly higher than 3). Note that an average value slightly higher than 3 still means that the low-GPA student still stated that he/she is neutral to or slightly interested in and liking science.
- GPA could explain ~80% of the variation in how well students liked and were interested in science.
- Students with a high GPA agreed that the CA3 course they took increased their liking of and interest in science, while students with a lower GPA found that the CA3 course did not increase their interest in science.
- Students with a high GPA had higher average self-efficacy increases than students with low GPA for statements 2, 4, 5, 7, and 12. Apparently, we are not effectively reaching and teaching science and technology to students with a low GPA.

5 FUTURE WORK

5.1 Fall 2010 workshop

A workshop will be held in fall 2010 as part of the dissemination phase of the spring 2010 work. The workshop will present the limited amount of data from the student-learning survey with a focus on an interactive discussion on how to more effectively carry out a future survey. The workshop should also discuss modifications to and/or clarifications of the existing CA3 learning goals.

Draft agenda:

- Presentation of results from further evaluation of data from 2008/2009 (very brief)
- Break-out session 1 - groups discuss interpretations of learning outcomes and comes up with examples of question types that could be used to assess each learning goal.
- Assembling suggestions from break-out session 1
    - Participation of committee member from CA3 learning goal committee

Lunch break

- Presentation of results from student learning assessment
- Break-out session 2 - groups discuss efficient ways of carrying out a future student learning assessment
- Assembling suggestions from break-out session 2

5.2 Assessing actual student learning

Guide lines for future assessments of student learning are expected to be one of the outcomes of the workshop to be held fall 2010.

5.3 Carrying out a revised student self-efficacy assessment

It was efficient to deliver the assessment on-line, so this format could be used in future assessments. However, many of the responses were difficult to evaluate. This was a result of unclear statements. Suggestions for revised statements were listed in the 2008/2009 assessment report.

Two additional improvements to a future self-efficacy assessment are suggested. The analysis carried out this year was based on both pre- and post-data delivered by individual students (paired data), where the number of pre and post-course data points necessarily is the same. Having both pre and post-course data allows calculation of improvements for individual students, which is not possible with un-paired pre- and post-course data. When the data was gathered in 2008/2009, students were given extra credit if they did the on-line survey. The first suggestion is that <u>students would only be given extra credit if they complete both the pre- and post-course assessment</u> in order to increase the number of paired data points in a future, revised self-efficacy assessment. Further, it was concluded last year that ~700 data points is sufficient to evaluate student self-efficacy assessments. However, this number of paired data points (read participating students) is not enough if we wish to sub-divide the data set to evaluate hypotheses based on additional criteria. The second suggestion is therefore to <u>maximize the number of participating students to increase statistical power.</u>