Learning Assistant Model for Teacher Preparation in Science and Technology (LA-TEST)

Project Annual Report, Academic Year 2007-2008

Project Activities & Findings

Edited by Benjamin Spike

Table of Contents

LA-TEST Project Overview and Research Questions.......................................................... 1

LA-TEST Project Publications and Talks........................................................................... 6

LA-TEST Project Year 2 Summary.................................................................................... 10

K-12 Team: Recruitment to K-12 Teaching ...................................................................... 15

K-12 Team: K-12 Teachers’ Classroom Practice .............................................................. 23

Conceptions of Teaching and Learning Team: The FASCI Instrument .............................. 28

Discipline-Based Educational Research Team: Summary .................................................. 53

Discipline-Based Educational Research Team: Physics...................................................... 55

Discipline-Based Educational Research Team: Astrophysical & Planetary Sciences ...... 69

Discipline-Based Educational Research Team: Applied Mathematics ............................. 74

Discipline-Based Educational Research Team: MCD Biology ........................................... 85

Discipline-Based Educational Research Team: Chemistry .............................................. 90
The Learning Assistant Model at the University of Colorado at Boulder

Project Overview

The Learning Assistant (LA) Model at the University of Colorado at Boulder uses the transformation of large-enrollment science courses as a mechanism for achieving three goals:

- To recruit and prepare talented science majors for careers in teaching
- To engage science faculty in the recruitment and preparation of future teachers
- To improve the quality science education for all undergraduates

The LA Program is a Teacher Recruitment Program, therefore students hired to be LAs typically have not already decided to become teachers.

One of the main purposes of the LA program is to find talented mathematics and science majors who might become interested in teaching as a result of their experiences working as an LA. Therefore, we hire a large number of LAs with the expectation that a fraction of them (typically 15%) will become interested in the complexities of teaching and decide to become K-12 teachers.

The transformation of large-enrollment courses involves creating environments in which students can interact with one another, engage in collaborative problem solving, and articulate and defend their ideas. To accomplish this, undergraduate LAs are hired to facilitate small group interaction in our large-enrollment courses at the same time they make up the pool from which we recruit new K-12 teachers.

The Difference

The differences between the LA model at the University of Colorado, Boulder and other standard models for undergraduate teaching assistants are:

- The focus on teacher recruitment and preparation
- A special seminar targeted at helping LAs integrate content, pedagogy, and practice
- A collaborative educational research program to evaluate the effects of the LA model
- The involvement of science research faculty in recruitment and preparation of teachers

Course Transformation and the Role of LAs

LAs are paid a modest stipend (currently $1,500/semester) to work approximately 10 hours per week in various aspects of course transformation. Examples of courses that have been transformed using LAs or another course reform are listed in Table 1.
Table 1. Transformed courses in AY 07-08. * indicates NEW for AY 07-08.

What is a Transformed Course?

There is no dictated design of what course transformation should look like. Instead, faculty members who are awarded LAs must agree to:

- use LAs to promote interaction and collaboration among students enrolled in the course
- meet in weekly planning sessions with the LAs who support their courses
- attend biweekly meetings with other faculty participating in the program
- attend a summer preparation session for using LAs
- evaluate transformations and assess learning in their own courses

Because there is little dictation as to what a transformed course should look like, there exist several models of course transformation among our participating departments. For example, one of two models of transformation in the physics department utilizes the University of Washington’s *Tutorials in Introductory Physics* in recitation sections each headed by one graduate TA and one undergraduate LA. The *Tutorials* involve conceptually-based group problem-solving activities which are based on research in physics education. LAs who work in *Tutorial sessions* formatively assess student understanding, ask guiding questions, and facilitate collaboration within groups. These tutorial sessions are supplemented by weekly lectures which are made interactive through infrared response systems and collaborative peer instruction.
A different model for course transformation is used in the Applied Mathematics department. In weekly LA-led problem-solving sessions, each small group of enrolled students uses a 2’x 3’ dry-erase board to collaboratively construct problem solutions.

A more radical form of transformation is found in the Astrophysical and Planetary Sciences department where one lecture per week is replaced by Learning Team sessions headed by LAs who facilitate collaborative problem solving as students analyze real astronomical data and generate and compare models to fit these data.

The LA Experience

Although the LA experience is somewhat different for each course, the experience for all LAs involves three related activities:

(1) Content: they meet weekly with their faculty instructor to plan for the upcoming week, reflect on the previous week, and analyze assessment data;
(2) Practice: LAs facilitate collaboration among learning teams by formatively assessing student understanding and asking guiding questions; and
(3) Pedagogy: LAs from all departments attend a special Mathematics and Science Education seminar where they reflect on their own teaching and learning and make connections to relevant education literature.

The Mathematics and Science Education Course

The Mathematics and Science Education seminar is jointly conducted by a faculty member from the School of Education and a K12 teacher. In this course, new LAs reflect on their own teaching practice, reflect on the transformations of the course in which they are working, investigate relevant educational literature, and engage in in-depth discussions about their own teaching and learning. Seminar readings and discussions include topics such as questioning techniques, learning theory, cooperative learning, student epistemologies, metacognition, argumentation, multiple intelligences, the nature of science, national standards, teaching with technology, and qualities of an effective teacher. Students
in this course try out new ideas each week in their learning teams and report their results in seminar. In many cases, LAs provide guidance to one another regarding managing issues that typically arise in their learning teams. Each week, LAs complete online reflections on their teaching and the learning of the students in their learning teams. In addition, throughout the semester LAs turn in two reflective essays that integrate the education literature with their own teaching and learning experiences. LAs often report that by studying and reflecting on student learning, they have become better learners themselves. At the end of each semester, LAs in the seminar present a poster session attended by their lead instructors, School of Education faculty, University of Colorado administrators, graduate students, and their peers. Each LA or small group of LAs present a poster that focuses on aspects of the LA experience that influenced their thinking both as a learner and as a teacher.

**Focus on Teacher Recruitment**

Although the LA experience is valuable for undergraduates who continue to any career, our program is specifically designed to actively recruit talented undergraduate students to careers in teaching. Therefore, a student can continue to be an LA for a second semester only if he or she shows commitment to finding out more about teaching. This may be evidenced by taking an education course or participating in an early K-12 field experience. LAs can be hired for a third semester only if they have been accepted to a teacher certification program at which time they are eligible for NSF funded Noyce Teaching Fellowships of up to $10,000 per year. As Noyce Teaching Fellows, students can become Lead LAs who mentor novice LAs, participate in the development course educational technology, or work with mathematics, science, and education faculty conducting educational research.
The LA-Test Project Research Questions

The Learning Assistant model for Teacher Education in Science and Technology (LA-Test) research project was designed to test the effectiveness of the LA model specifically in terms of LAs’ development of content knowledge, pedagogical knowledge, and their practice in K-12 schools. Faculty members from education, mathematics, and science, K-12 teachers, graduate students, and Noyce Fellows comprise three interacting research teams: the Discipline-Based Educational Research (DBER) team, the Conceptions of Teaching and Learning (CTL) team, and the K-12 team. These interacting research teams investigate teacher recruitment rates as well as the research questions shown in table 2 and synthesize results on an ongoing basis.

Each research team focuses on a specific set of questions (shown below). Answers to these questions will be synthesized by the end of the project in order to make inferences about the impact of the Learning Assistant (LA) Model on teachers’ pedagogical content knowledge.

The DBER team consists of graduate students from five departments: Physics, Chemistry, Applied Mathematics, Astrophysical and Planetary Sciences, and Molecular, Cellular, and Developmental Biology. The CTL team consists of faculty members from the School of Education in Research and Evaluation Methodology and from the Physics department, and graduate students in Science Education and in Physics Education Research. The K12 team consists of faculty members and graduate students in Science and Mathematics Education.

Research Questions:

DBER Team:
(a) How do LAs compare to other STEM majors in terms of their content understanding, beliefs about the discipline, and beliefs about learning in the discipline?
(b) What effects can be observed on student achievement in courses that are supported by LAs?

CTL Team:
(a) What is the effect of the LA model on the sophistication of LA pedagogical understanding?
(b) Does sophistication of pedagogical understanding vary by length of exposure to the LA model?
(c) How is the pedagogical sophistication of STEM LAs different from the sophistication of STEM non-LAs who become teachers?

K-12 Team:
How do teachers and teacher candidates who participated as LAs compare to those who did not in terms of:
(a) Practicum-based coursework
(b) Their teaching practices
(c) K-12 student attitudes and beliefs about mathematics and science
(d) Retention and attrition rate
AY 2007-2008 Talks, Workshops, Publications

Publications

Workshops
3. Workshop conducted for the national Physics Teachers Education Coalition (PTEC) and funded by the American Physical Society, held at University of Colorado, Boulder.

Talks
8. N.D. Finkelstein, "Understanding the Tools Students Use to Learn Physics: Representation, Analogy, and Computer Sims" Joint Colloquium Center for


11. S. Pollock, N. Finkelstein, L. Kost*, “Understanding the Gender Gap in Introductory Physics” Session W16, APS meeting, St. Louis MO, Apr 15 2008


19. C. Turpen*, N. Finkelstein, "Not all Interactive Engagement is the Same: Variation in Faculty Use of Peer Instruction" AAPT, Greensboro, NC, Aug 2007.
27. Mary Nelson, Invited talk to Presidential Teaching and Learning Scholars, Spring Retreat, March 2008
32. Otero, V. (2007, August). The Road Less Traveled: The STEM Colorado Learning Assistant Program. Presented at the bi-annual meeting for the American Association of Physics Teachers, Greensboro, N.C.
LA-TEST Project Summary Findings

K-12 Research Team Summary
Valerie Otero

There were two main activities in which the K-12 research engaged: studying the LA Model’s effect on recruiting new students to careers in teaching in comparison to other early teaching experiences and studying former LAs and Noyce fellows currently teaching full time in the districts in terms of their practice. Former LAs are compared to other early career K-12 teachers who graduated from the same certification program but did not have the LA “treatment.” Both of these K-12 research projects are summarized below.

Laura Moin is a faculty member in the Science Education research group at CU-Boulder. Over the past three years she has been conducting a controlled research study to determine the true effect of the Learning Assistant/Noyce fellowship program for recruiting new teachers. She has been collecting survey data on students’ interests in teaching from the Learning Assistant recruitment/information sessions that are held mid-semester each semester. She then follows the students that do not apply to be LAs, those that apply to be LAs and are not hired, and those that apply to be LAs and are not hired, and continues to survey these students to find out whether their interests in teaching increase over a semester, a year, and beyond. Some preliminary data from Prof. Moin’s research are shown in figure 1 below. As is evident on the graph, only those who participated in the LA program showed an increase in interest in teaching. For more details, please see the section on K-12 recruitment.

Figure 2. Interest in teaching. The leftmost bar shows students who did not seek out new teaching experiences after applying but not being accepted to the LA program, the middle bar shows students who did seek out and engage in other teaching experiences after not serving as LAs, and the right-most bar shows LAs interest in teaching increase over a year.
David Webb heads the K-12 research team associated with the LA-TEST research project. Recently Professor Jeff Frykholm has joined this team. A team of graduate students is hired each semester to do observations in former Noyce Fellow’s classrooms. A control group of students who were certified through the same program but did not have the Noyce/LA “treatment” is also observed. Each teacher is observed three times throughout a year and each observation consists of two researchers evaluating teaching using the same instrument (the Reformed Teacher Observation Protocol). At the end of each observation, the two researchers reconcile scores. All of the graduate student researchers who go out into the field and make observations attended over 10 hours of training using the instrument and the entire group reached a high reliability in their scoring in preliminary observations. Some preliminary results are shown in Table 1. The numbers are still small for this study mainly due to the amount of time it takes to get full school and district approval for work in the individual K-12 classrooms. We expect greater numbers and more types of data in the upcoming year. Other data that this group is collecting include scores on attitudinal measures for middle and high school students of these (and other) teachers and artifact packages representing an entire unit of instruction which we have already begun collecting from each of the teachers studied. In addition, each teacher who participates in the study is interviewed three times throughout the year. A coding scheme has been established and all data are currently being coded. We feel that these data represent an objective measure of the program especially since the data are collected by graduate students who have no contact with the study group prior to the data collection associated with the LA-TEST project.

Table 1. Reformed Teacher Observation Protocol Data from LA-TEST project

<table>
<thead>
<tr>
<th>Experience</th>
<th>Math</th>
<th>Science</th>
<th>Overall Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA (N=8)</td>
<td>68.3</td>
<td>42.1</td>
<td>49.6</td>
</tr>
<tr>
<td>Non-LA (N=4)</td>
<td>36.0</td>
<td>41.5</td>
<td>39.6</td>
</tr>
</tbody>
</table>

In summary, our approach to evaluation combines the necessary objective measures of success with an access to rich and relevant data, by leveraging research that is currently being conducted in synergist studies. The Principal Investigators of these studies are faculty at CU, but not directing the Noyce program and provide a sense of outside evaluation. We did not discuss the DBER findings here, but these too are collected electronically and analyzed using standard non-subjective methods that we feel are objective measures of students’ content knowledge (these data were discussed in detail in our proposal). We initially felt that an affordable external evaluation would not provide the richness and relevance of data and analyses that are currently a part of our ongoing LA-TEST research project. However, although we did not originally budget for an external evaluator, we could allocate a 25% to 50% graduate student line from the LA-TEST project to fund a graduate student to act as an external evaluator of the Noyce project. We did this for our first year of the STEM-TP project, and these two 25%
graduate students maintained distance from the project and objectivity in the data analyses. These students would come from the Research Evaluation Methodology program and their work would also be consistent with the LA-TEST project goals and design. We would prefer to use these LA-TEST funds for the type of research described above, however we would be willing to launch an external evaluation of the project if you feel that it is necessary.

We are not comfortable using the Noyce funds for the evaluation because the fraction allocated for non-scholarship activity is only enough to run and expand the project according to our Phase II goals. In addition to funding a part-time administrative assistant and faculty release time for tracking and working with Noyce Fellows, we hoped to use some of these funds for mini-grants for Noyce Fellows as they move out into the field and become teachers.

**Conceptions of Teaching and Learning (CTL) Team**

**FASCI Instrument Summary**

Robert Talbot and Derek Briggs

The Conceptions of Teaching and Learning (CTL) Research Team is working on the development and validation of the Flexible Application of Student-Centered Instruction (FASCI) survey instrument. The research team consists of Professor Derek Briggs and doctoral candidate Robert Talbot, with assistance from Professor Valerie Otero.

During the academic year 2007-2008, a pilot test of the FASCI was conducted in which 65 individuals responded to the open-ended items on the survey. Item and scoring rules were further refined before and after this pilot test and response data was analyzed both qualitatively and quantitatively.

We have found that our FASCI items span a range on difficulty on both the Flexible Application (FA) and Student-Centered Instruction (SCI) dimensions of the instrument. Scores on the FA dimension are more reliable than those on the SCI dimension. We believe that scores from the current version of the instrument can be used to make qualitative distinctions among respondents with respect to their FA sophistication, such distinctions can be made less reliably with respect to their SCI sophistication. Preliminary evidence in support of the validity of FASCI scores indicates that students participating in the Learning Assistant program tend to score higher on both dimensions than non-LA students enrolled in the School of Education teacher certification program.

**Future Directions**

One future direction for our work will focus on the potential threat to validity of score interpretations posed by the content-neutrality of the current FASCI items. We will be developing and administering a physics-specific version of the FASCI in order to investigate this. We will also be working with the LATEST K-12 team to observe the actual practice of teachers and compare that data to FASCI scores from those teachers in a further effort to establish a validity argument for the FASCI. Finally, we will be developing new FASCI items and revising our scoring rules in the attempt to increase the reliability of SCI scores.
The primary research questions for the DBER team concern the performance of students and student educators in LA-supported science courses:

1. *How do LAs compare to other STEM majors in terms of their content understanding, beliefs about the discipline, and beliefs about learning in the discipline?*

Since LA selection is based partially on performance in lower-division courses, they tend to enter with comparatively high conceptual and belief survey scores. Nonetheless, the DBER team demonstrates continued learning gains and positive belief shifts as a result of the LA experience. Departments that examine LA conceptual knowledge report that some LAs have even been observed to out-perform graduate TAs on conceptual surveys. Longitudinal studies have also shown that students who were LAs tend to perform better than non-LAs in upper-division coursework.

The DBER team has also demonstrated increased LA interest in education, through survey data and in the form of participation in volunteer-based K-12 outreach programs. Some have even chosen to return to the LA program on a volunteer basis. This observation has, in part, inspired the creation of the CU Teach program, in order to encourage and recruit STEM students with an interest in science education.

The primary tool for measuring educational and disciplinary beliefs is the CLASS survey, which began as a physics survey and has been adapted to suit various courses. The use of conceptual surveys is less uniform, however; some departments continue to use established conceptual surveys, while others have not developed one. Standardizing conceptual evaluation across the DBER disciplines remains a challenge due to the great diversity of subject matter.

2. *What effects can be observed on student achievement in courses that are supported by LAs?*

The DBER team continues to observe measurable, positive effects on students in reformed courses during AY 07-08. These effects include improved learning gains, decreased failure rates, and greater student satisfaction with coursework.

In some cases, the recent adoption of LAs affords no measurable change compared to previous semesters, and tools are being developed to more carefully observe the effects. Negative shifts in beliefs have also been observed to carry over from traditional to reformed courses, which remains an active subject of inquiry.

Examples of effects observed across the DBER team for AY 07-08:
• In physics, normalized learning gains continue to be higher than in traditionally taught courses. The effect appears to be dependent on the instructor’s familiarity with PER methods and reforms.

• Significant drop in failure rate during AY 07-08 as a result of reforms in Calculus I & II. In Spring 08 the failure rate of Calc II fell to 17% for the first time, from 31% and 27% in previous years.

• MCDB students who experienced weekly interactions with LAs had significantly higher learning gains and post test scores, and performed better in the class overall than their peers.

• More positive student course rating in APS, with specific mention of help provided by LAs.

Departments have continued to expand the number of LA-supported courses, and the number of professors requesting LA support increases with each academic year. In addition, all departments report increased departmental support in the form of funding, building space, and instructors. The favorable results observed in lower-division courses have also prompted departments to examine teaching methods in upper-division courses, as was done with PHYS 3310 in AY 07-08. The findings strongly suggest that the effects of these reforms are not limited to introductory coursework.

Each project is described in detail throughout the remainder of this document.
K-12 Team

Recruitment to K-12 Teaching Careers

Laura Moin

Activities

Data Collected

- We distributed surveys during information sessions and in the summer. The surveys aim at conducting program evaluation in a comparative perspective: what is the effect of the program (changes in teaching career interest and in views about teaching) on program participants in comparison to individuals not participating in the program (seeking other teaching experiences or not seeking any teaching experience)?
- About 150 information session surveys (annually) and 250 summer surveys were collected.
- Data collection is ongoing, every information session and summer.

Findings

Information from information session surveys and application packages indicate that 60% of information session attendees applied to the program. From the information session surveys, it seems that applicants and non-applicants did not significantly differ in their views about a teaching career neither did they differ in their K-12 math/science teaching career interest (although among non-applicants there was more variability in this variable). Applicants and non-applicants did significantly differ, however, in that more applicants tended to see the program as a possibility to gain a teaching experience while more non-applicants tended to seek in the program an opportunity to earn money.

However, the online survey revealed that half of non-applicants (and most admitted applicants who declined the offer) had scheduling conflicts that prevented them from applying to the program and there were no non-applicants indicating low pay as a reason for not applying. Hence, while the group of applicants may be more homogeneous in terms of their teaching career interest, it cannot be said that the group of non-applicants is not interested. This last group is more heterogeneous with respect to their teaching career interest, some non-applicants are just as interested as applicants but cannot yet apply for various reasons. Other non-applicants are not as interested in a teaching career.

From the 60% of information session attendees who applied to the program 57% of them were not admitted. Given the considerations above, the remaining 40% of non-applicants may not necessarily be less interested in a teaching career; however, only few of them sought other teaching experiences after their contact with the program (2 out of 47). The same is true for non-admitted applicants (8 out of 103 sought other teaching experiences after attending the information session) and for admitted applicants who had to decline the offer (2 out of 9 admitted sought other teaching experiences after their contact.
with the program). It seems that there is a momentum of teaching career interest when they contact the program and many of them apply. After that momentum, interest in a teaching career significantly declined for all those who did not seek a teaching experience (n = 64, t = -2.3, p < 0.05) and even those who sought a teaching experience outside of the program (mostly tutoring) (n = 14, t = -2.7, p < 0.05).

Before participating in the program, non-applicants, non-admitted applicants, and admitted applicants only differed in two appealing teaching factors. More admitted applicants than non-admitted and non-applicants indicated the affective exchange involved in teaching as an appealing factor (F = 6.4, p < .01). Also, non-applicants significantly differed from admitted applicants in the available teaching jobs as an appealing factor for a teaching career (F = 4.3, p < .05). After participating in the program, the affective exchange involved in teaching became only marginally significant. Participation in the program affected the view of teaching as a complex task; significantly more program participants indicated the complexities of teaching as an appealing factor as compared to non-admitted applicants and non-applicants (28%, 12% and 14% respectively) (F = 3.5, p < .05). Also, the view of teaching as a creative job changed upon participation in the program; non-admitted applicants were significantly more likely to express this view as an appealing factor than were admitted applicants and non-applicants (29%, 17%, and 12% respectively) (F = 3.4, p < .05).

Hence, it seems that self-selection to apply is driven by a combination of available time to commit to the program, not much interest in seeking a teaching experience but an interest in job opportunities and earning money. Applicants who view teaching as a task that involves affective exchange (which may subconsciously be transmitted in the application personal statement or the interview) were more likely to be admitted although this view did not persist after participating in the program. It seems that program participants gained an understanding of the complexities involved in teaching. We are uncertain as to why upon participation in the program, teaching tended to be viewed as less creative than before participating in the program. It might be related to gaining an understanding of the theories and principles that support good teaching (learned in the LA seminar) that participants might consider opposite to free creativity.

In addition to controlling for self-selection to participate in the program, we also controlled for trend in the population. We longitudinally tracked students’ answers to questions, in particular whether they had sought another teaching experience and their involving interest in teaching careers. Among this data set, we identified 110 respondents who participated in the program and answered one instrument before and another instrument after program participation; 63 respondents to two instruments, including applicants and non-applicants, who did not seek any new teaching experience between the two surveys (they may have had a teaching experience from before attending the information session), and only 14 respondents who sought a new teaching experience outside of the program after attending the information session. We conducted a one-way ANOVA on change in teaching interest among the three groups of respondents. The changes in K-12 math/science teaching career interest were statistically significant (F = 12.8, p < .01). Post-hoc comparisons showed that program participants significantly gained interest as compared to those who did not seek any teaching experience and to others who sought other teaching experiences (mostly tutoring).
Who applies to the program? Comparison of applicants and non-applicants

We asked new information session attendees for their interest in a teaching career using the scale: 0 = I am not interested in becoming a K-12 math/science teacher, 1 = I am curious about K-12 math/science teaching but not as a career, 2 = K-12 math/science teaching is one of my many career options, 3 = K-12 math/science teaching is among my top career options, and 4 = I want to become a K-12 math/science teacher. The distributions of teaching career interest for applicants (mean=1.84, s = 0.98, n=291) and for non-applicants (mean=1.67, s = 1.08, n = 198) are shown in Figure 1 below. While the means of both groups did not significantly differ (t = -1.75, p > .05), their standard deviations did (F = 5.33, p < .05) indicating a wider spread for non-applicants.

![Figure 1. Applicants and non-applicants new-attendee’s interest in K-12 teaching](image)

Information session survey respondents also indicated up to 2 reasons why they attended the information session. The options were: a) I had an LA in one of my courses and thought that what s/he did was cool, b) Somebody told me about being an LA and I thought it’d be cool, c) I was told I’d be good at being an LA, e) Seeking a teaching experience e) I am looking for an opportunity to earn some money, f) I am looking for an opportunity to get more involved in my discipline (for example to learn the content more deeply or interact closely with faculty), g) My friend was coming and I just came along, h) Free food, i) Other, specify. Responses are shown in Figure 2 below. “Other” reasons expressed by new attendees included: “Extra-curricular activity,” “Teaching is one way of understanding the material better,” “1st astronomy class and I really enjoy it,” “My sister [name] told me it was a great learning experience to help others,” “I came to see [name],” “I like tutoring and want to use this as a resume builder for med school,” “It is a good work...
experience on my resume,” “Improve my odds for grad school,” “My [illegible] professor told me it would help develop better teaching techniques.”

There were no significant differences in the proportion of applicants and non-applicants indicating various reasons for attending the information sessions, except for the reason seeking a teaching experience. As expected, significantly more applicants than non-applicants indicated this reason to attend the information session (t = 2.30, p < .05). Seeking a teaching experience was also statistically significantly correlated with teaching career interest (r = 0.26, p < .00). We also found a statistically significant correlation, but of a different sign, between teaching career interest and seeking an opportunity to earn money (r = -0.15, p < .01) but the proportions of applicants and non-applicants indicating this reason did not significantly differ.

![Figure 2: Reasons for attending the information sessions](image)

Information session new attendees expressed their strongest motivators and de-motivators for teaching careers. The proportions of applicants and non-applicants who indicated the main motivators and de-motivators for teaching careers did not significantly differ. The choices for motivators were: a) having an impact on people’s lives, b) summer vacations and school schedule, c) not a highly demanding job (in terms of time or intellectual demands), d) the complexity of teaching, e) being an authority figure, f) being a
role model, g) the affective exchange with children, h) the low risk nature of the job (mistakes wouldn't kill anybody or send anybody to jail, or bankruptcy, etc.), i) the number of job opportunities, j) the possibilities of further education and professional development, k) creativity in teaching, l) high prestige of teachers in society, and m) others with specification. The strongest motivators were the impact of teaching on people’s lives (indicated by 78% of new attendees), the creativity involved in teaching (indicated by 28% of new attendees), the affective exchange (indicated by 25% of new attendees). Professional development opportunities and being a role model were each indicated by 20% of new respondents. School schedule and summer vacations, lower scholarly demands as compared to other professions, high prestige of teaching, and available job opportunities were each indicated by less than 20% of new attendees.

The choice for de-motivators to teaching careers were: a) routinized work, b) influencing somebody’s life, c) no need of further education, d) lower scholarly demands as compared to higher academic institutions, e) limited possibilities of professional advancement to more powerful positions, f) low salaries, g) no free choice of what to teach or what to do, h) dealing with parents and administrators, i) dealing with students behavior, discipline, and classroom management, j) being in front of a class, k) job not highly regarded in society, and l) other. The strongest de-motivator for teaching careers was low salary (indicated by 43% of new attendees), followed by dealing with parents and administrators (indicated by 30% of new attendees) and dealing with students’ discipline and behavior (indicated by 29% of new attendees). Routine work was a de-motivator for 25% of new attendees, and low prestige, low scholarly demands as compared to other professions, and limited possibilities of career advancement were each indicated by less than 20% of new attendees.

In sum, from the information session surveys, it seems that for the most part, applicants and non-applicants did not significantly differ in their views about a teaching career. Applicants and non-applicants did significantly differ, however, in that more applicants tended to see the program as a possibility to gain a teaching experience while more non-applicants tended to seek in the program an opportunity to earn money. Maybe, they did not apply because the pay was not high enough for them? This is a question for which we sought an answer in the online survey. A paradox also emerged from the information session survey study. While more applicants indicated that they sought a teaching experience and all those who indicated that they sought a teaching experience also tended to express a higher teaching career interest, applicants and non-applicants did not significantly differ in their the mean for teaching career interest. A possible explanation will also be found in the reasons for not applying from the online survey.

Non-program participants – Online survey

The online survey was distributed to 5 groups of respondents: prior semester program participants, admitted applicants for next fall semester, and three sub-groups of non-program participants: non-applicants, admitted applicants who declined the offer, and non-admitted applicants.

Non-applicants. Findings from non-applicants should be regarded with caution because response rate for this group on the online survey was barely above 23%. Table 3 below shows the reasons for not applying and whether they are planning on applying in the
future. Other reasons for not applying including putting it off for later time in college, being abroad, not feeling academically prepared, and not having been aware of the program or opportunity to apply. Although earning money was a reason to attend the information session more frequently cited by on-applicants, no respondent indicated low pay as a reason for not applying.

Table 3: Respondents by intention to apply in the future and reason for not yet applying

<table>
<thead>
<tr>
<th>Reasons for not yet applying</th>
<th>Planning on applying in the future</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Specific scheduling conflicts or overloaded schedule</td>
<td>4</td>
</tr>
<tr>
<td>I didn’t think I would be accepted due to low grades</td>
<td>2</td>
</tr>
<tr>
<td>No positions of my interest</td>
<td>0</td>
</tr>
<tr>
<td>I missed the deadline</td>
<td>1</td>
</tr>
<tr>
<td>I realized that teaching is not what I want</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>

From the 60% of information session attendees who applied to the program 57% of them were not admitted. Given the considerations above, the remaining 40% of non-applicants may not necessarily be less interested in a teaching career; however, only few of them sought other teaching experiences after their contact with the program (2 out of 47). The same is true for non-admitted applicants (8 out of 103 sought other teaching experiences after attending the information session) and for admitted applicants who had to decline the offer (2 out of 9 admitted sought other teaching experiences after their contact with the program). It seems that there is a momentum of teaching career interest when they contact the program and many of them apply. After that momentum, interest in a teaching career significantly declined for all those who did not seek a teaching experience ($n = 64$, $t = -2.3$, $p < 0.05$) and even those who sought a teaching experience outside of the program (mostly tutoring) ($n = 14$, $t = -2.7$, $p < 0.05$).

Non-applicants and non-admitted applicants did not significantly differ in their K-12 math/science teaching career interest, gender, academic year distribution, academic performance (GPA) or teaching experience. But they did significantly differ in their view of teaching as a creative job. Only 12% of non-applicants indicated the creativity of teaching among the two most appealing aspects of teaching but about 29% of non-admitted applicants indicated the creativity of teaching as an appealing factor in teaching. Hence, self-selection to apply may respond to a view of teaching as a creative job (if there are no scheduling conflicts).

What happened to those who self-selected to apply? After self-selection to apply, there is program-selection to participate. In general, instructors’ criteria to admit participants to the program include: (1) Interest in teaching (at any level); (2) Personable and able to work well with others; and (3) High performance on the class/discipline in which they want to experience teaching. Among the 474 applicants in the six semesters,
170 were admitted to the program and 304 were not\footnote{The number of applications is considerably larger than the number of information session surveys because a few applicants did not attend the information session and mostly because the last information session was attended by an unexpectedly large number of attendees and the printed survey were not enough. The surveys were randomly distributed at this last information.}. We found a significant difference in disciplinary GPA between admitted and non-admitted applicants, 3.6 compared to 3.2 respectively (p < .01).

Instructors also made their selections of applicants by department and many departments interviewed their applicants. We found three other significant differences between program participants (after participation) and non-admitted applicants. First, program participants indicated a significantly higher interest in a K-12 math/science teaching career (2.0 versus 1.5, p < 0.01). Second, a lower proportion of program participants indicated summer vacation and school schedule as an appealing aspect of teaching (6% versus 19%, p < .01). Finally, a higher proportion of program participants the complexities of teaching as an appealing aspect of teaching (28% versus 12%, p < .05). Are these the effects of program participation? Maybe. However, to more reliably attest to the effect of program participation, it is necessary to compare the change in program participants’ K-12 math/science teaching interest compared to control groups of respondents who may or may have not sought other teaching experiences.

In order to do so, we longitudinally traced 162 respondents who responded to either two information session surveys or one information session survey and the on-line survey. We selected this two-survey response criterion because it provided a means for longitudinally tracking students’ answers to questions, in particular whether they had sought another teaching experience. This will help us control for trends existing in the population and which are independent from program effects. Among this data set, we identified 51 respondents who participated in the program and answered one instrument before and another instrument after program participation; 32 respondents to two instruments, including applicants and non-applicants, who did not seek any new teaching experience between the two surveys (they may have had a teaching experience from before attending the information session), and only 7 respondents who sought a new teaching experience outside of the program after attending their first information session. We conducted a one-way ANOVA on change in teaching interest among the three groups of respondents. The changes in K-12 math/science teaching career interest were statistically significant (p < .01). Post-hoc comparisons showed that program participants significantly gained interest as compared to those who did not seek any teaching experience and to others who sought a wide variety of teaching experiences. These findings are shown in Figure 6 below.
Respondents who sought a teaching experience outside of the program mostly did tutoring. Tutoring as a teaching experience is radically different than the college teaching experience offered by the program and hence the effect of the program may be due to the type of teaching experience offered rather than the program as a whole.

**Future Directions and Anticipated Year 3 Goals**

- We are finding answers to the main research questions: a) Who are the STEM undergraduates more likely to be attracted to a teaching career; b) Why are some STEM undergraduates not attracted to teaching; c) How to attract them. In addition, we are interested in conducting more valid and reliable program evaluation in a comparative perspective of the effects of recruitment program participation versus non-participation and seeking other teaching experiences.
- The same stuff will continue this project.
- Future directions for the project:
  1. Continue to collect data to investigate the effect of the type of teaching experience (college teaching) as oppose to other teaching experiences. It may be that the positive teacher recruitment effect we found for the LA program is due to the type of teaching involved rather than other components of the program (for example the seminar).
  2. More data will also help answer still unexplained findings (for example why program participants tended to see teaching as less creative).
  3. In addition, as another recruitment program will take place at CU, comparisons between programs will be mostly enlightening to identify factors that prompt (or not) candidates to pursue a teaching career.
K-12 Team

K-12 Teachers’ Classroom Practice

David Webb, Kim Geil, Steve Iona, Mary Nelson, Kara Gray, Michael Ross, Jeff Frykholm

Activities

The overarching goal of the LA-Test K-12 research team is to conduct a quasi-experimental study comparing the beliefs and practices of Learning Assistants who become secondary teachers with teachers of comparable experience who completed a CU-Boulder licensure program with no experience as an LA. Under the purview of the LA-Test research project, our approach and method is designed in particular to study the effects of the LA program among practicing K-12 teachers and their students. As a post-undergraduate follow-up study, our research involves interview, observation of practice \textit{in situ}, and collection of classroom artifacts across various school contexts to provide objective portraits of teacher beliefs and practice.

The LA-Test K-12 research team is led by Dr. Webb, and includes two other faculty (Dr. Nelson and Dr. Frykholm), three doctoral students (Geil, Gray, and Ross), and a retired science educator (Iona). The subject background of the research team is evenly split between mathematics education (Webb, Frykholm & Nelson) and science education (Iona, Gray, & Ross); and Kim Geil’s research focus is educational psychology.

Data Collected

In fall 2007, we identified LA teacher candidates who had found teaching positions in math and science and also identified potential non-LA (control) teachers who had graduated in the same year. We were able to secure participation from eight former LA students and four non-LA students, split evenly among math and science candidates (see Table 1).

\begin{table}[h!]
\centering
\begin{tabular}{|c|c|c|}
\hline
 & Science & Math \\
\hline
LA & 4 & 4 \\
Non-LA (control) & 2 & 2 \\
\hline
\end{tabular}
\caption{Year 1 K-12 cohort experience and content area}
\end{table}

Of the 11 former LA students that we contacted, 2 declined to participate because they were not currently teaching and the other did not return phone calls or email. Of the 13 non-LA students we contacted, 6 did not respond to multiple invitations to participate, and 3 declined to participate. The three control teachers who replied to our invitation, declined to participate due to concerns about committing to a research study during their first year of teaching. District approval for current and future teacher participation was secured for each participant in addition to permission from school principals to conduct research at the school site.
Since participation in this study was voluntary, we assume that the difficulty in recruiting or eliciting replies from potential control teachers might have been due to a lesser commitment to research in teaching and learning among non-LAs. This is in contrast to the LA students who were immersed in such studies during their undergraduate experience and had ongoing contact with math, science and education faculty over several years.

Given the complexity of classroom-based research and the many variables that could be examined, we first developed a coding scheme that included the a set of primary variables that could be examined across all data sources. These LA Test K-12 Constructs of Interest (LAT-CoI) are:

1. Assessment practices (methods and conceptions)
2. School Context (perception of support from colleagues and administrators)
3. Construction of knowledge (i.e., teachers conceptions of how students learn math/science)
4. Teacher background (including LA and teacher education experiences)
5. Teacher self-efficacy (i.e., perceptions of and responses to problems of practice)
6. Views of students (how they regard student diversity, readiness to learn, students’ struggles)
7. Lesson Design/Planning (regard for student prior knowledge, inquiry, content)

Secondary and tertiary codes have also been developed to elaborate different aspects of these constructs of interest, exemplified in the various data sources. Since the research questions and related coding scheme call for an examination of teachers’ conceptions and practices, the data sources selected drew from a blend of interview protocols, survey instruments, classroom observation protocols and sampling of classroom artifacts. The data sources used in Year 1 of this study are as follows:

Table 2. LA-Test K-12 research: Year 1 data sources

<table>
<thead>
<tr>
<th>Teacher Conceptions</th>
<th>Teacher baseline interview</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observation interview</td>
</tr>
<tr>
<td></td>
<td>Reformed Teaching Observation Protocol</td>
</tr>
<tr>
<td>Teacher Practice</td>
<td>Scoop notebook</td>
</tr>
<tr>
<td></td>
<td>Reformed Teaching Observation Protocol</td>
</tr>
<tr>
<td>Students’ Conceptions</td>
<td>Colorado Learning Attitudes Science Survey (adapted for secondary school use)</td>
</tr>
</tbody>
</table>

Baseline interviews were conducted with all teachers upon their entry into the study. Three interviews were conducted over the phone and the rest were conducted in person, at their school site. Three classroom observations were conducted for 10
participants, which included use of Reformed Teaching Observation Protocol (RTOP) and the observation interview.\(^2\)

To support the collection of classroom artifacts, we video-taped and processed DVDs that included an overview of Scoop notebook procedures for teachers, to provide an additional resource for teachers for future reference (and for out-of-state participants). Each Scoop notebook includes teacher lesson plans, assessments, assignments, and reflections about lessons over a period of one week. Scoop notebooks were collected from 11 of the 12 participating teachers. We also plan on reviewing the field notes and artifacts embedded in the RTOP for evidence of teachers’ conceptions of inquiry and student learning.

The Colorado Learning Attitudes Science Survey (CLASS), which had been used with LAs during their undergraduate experience, was reformatted so that it could be administered in paper-and-pencil format. Math and science versions of the CLASS were developed, and the language was also adapted so that it could be used with middle grades and high school students. Spanish translations of the math and science CLASS were also developed. Eleven participating teachers administered the CLASS to one class of their students.

**Findings**

The coding of transcripts for baseline interviews is 75% complete and 20% of the observation interviews have been coded. By mid-August, the rest of the baseline and observation interviews will have been coded and prepped for use in case study reports for each teacher.

All RTOP data has been translated into an Excel file and summarized as descriptive statistical data for the LA and non-LA groups. Using the standard protocol for analysis, preliminary analysis of RTOP data suggest that the LA group (N[obs]=21, mean = 49.62, std dev = 21.4) is more likely to engage in inquiry-oriented classroom practices than the control group (N[obs] = 9, mean = 39.67, std dev = 14.7). However, it is worth noting that these differences, while large, are not statistically significant comparisons when contrasting the 95% confidence intervals for each group’s mean. The relatively low numbers of participants in each group, especially the control group, limit statistically significant results. Although the difference between LA and non-LA groups can be largely attributed to differences in ratings among mathematics teachers, even the large differences between means for LA and non-LA mathematics are not statistically significant due to the relatively low number of participating teachers. The additional participants we expect to add to the study next year will likely produce more meaningful, and perhaps statistically significant, comparisons.

---

\(^2\) Two math participants (one LA, one non-LA) had teaching assignments in South Dakota and Vermont. The agreed to participate in all phases of the data collection with the exception of classroom observation.
Table 3. RTOP overall mean (std dev): Teachers w/ LA and w/o LA experiences by content area

<table>
<thead>
<tr>
<th>Experience</th>
<th>Math</th>
<th>Sci</th>
<th>Overall Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA</td>
<td>68.3 (20.1)</td>
<td>42.1 (13.9)</td>
<td>49.6 (19.6)</td>
</tr>
<tr>
<td>non-LA</td>
<td>36.0 (29.5)</td>
<td>41.5 (17.7)</td>
<td>39.7 (20.5)</td>
</tr>
<tr>
<td>Overall Mean</td>
<td>57.6 (27.0)</td>
<td>42.0 (14.6)</td>
<td>46.6 (20.1)</td>
</tr>
</tbody>
</table>

Figure 2 highlights differences in mean RTOP sub-scores for lesson design and implementation, lesson content, and classroom culture. Ratings for each observation prompt in the RTOP are marked on a scale from 0 to 4. As shown in the bar graph, the overall mean scores for teachers with LA experience are higher than the overall mean scores for teachers without LA experiences. These differences, however, are not statistically significant but are, perhaps, suggestive of differences we might expect to observe among participating teachers in subsequent years.

Figure 1. RTOP sub-score means (std dev) for teachers with LA and non-LA experiences

With respect to the Scoop notebooks, Dr. Frykholm is leading the analysis of this data source and recently constructed an analysis table to facilitate cross-referencing to the constructs of interest. Two notebooks have already been analyzed and we expect to complete the analysis of all Scoop notebooks by early fall 2008.

To date, all paper and pencil data for CLASS have been transcribed and summarized as an Excel file. The response rates for CLASS in Year 1 were disappointingly low. We assume this result was due to low student attendance in some teachers’ classrooms and difficulty in securing consent from some parents for their student to participate. In short, the Year 1 response rates from the CLASS are problematic and may preclude any
reasonable analysis or comparison between groups. Consequently, in Year 2 of this study we have decided to administer the CLASS to at least three classes for each teacher to increase the number of student responses per teacher.

**Conference papers and publications**

To date, there have been no conference presentations or manuscripts developed that related to the LA-Test K-12 research. As the treatment and control groups increase next year, we expect to have findings that will be of interest to the research community and expect to submit proposals for conference papers and at least one manuscript for publication.

**Future directions**

Next year we plan to use the same approach to data collection with some minor modifications to ensure greater participation among potential non-LA teachers and administration of the CLASS to secondary students. The additional number of teachers available in next year’s cohort is also quite large, and so we expect to recruit another 20 teachers in addition to the 12 teachers who participated in the study over this past year.
Conceptions of Teaching and Learning (CTL) Team

FASCI Instrument

Robert Talbot and Derek Briggs

Activities

Personnel

Derek Briggs, an assistant professor in the School of Education at CU-Boulder has led the activities of the CTL team in developing and validating the FASCI instrument. Dr. Briggs is chair of the Research and Evaluation Methodology program and specializes in psychometrics and educational measurement.

Robert Talbot, a fourth-year doctoral student in science education, worked half-time on the development, administration, and analysis of FASCI data during the 2007-2008 academic year. He is a former high school science teacher and has experience as a teacher educator and a background in educational measurement. Two other doctoral students assisted in the scoring of FASCI response data on a part-time basis.

Sample

In the 2007-2008 pilot test of the FASCI instrument, 65 individuals responded to the 5 scenario-based tasks on the current version of the instrument. Of these 65, 28 were female and 37 were male. The majority (n = 43) of respondents were undergraduate Learning Assistants (LAs) at our university. Two respondents in the sample were Noyce Fellows within the Learning Assistant program (Otero, 2006). Faculty Experts from the School of Education and the STEM departments were also included in the sample (n = 3). It should be noted that these faculty are known as exemplary teachers and committed educational researchers and therefore are not typical faculty. Graduate students from STEM departments who were teaching assistants for undergraduate courses and labs were also included in the pilot sample (n = 8). A sample of practicing K-12 science teachers from local area schools (n = 8) was also surveyed. One respondent did not disclose their group affiliation. Mean self-reported teaching experience for the entire sample was 4.5 years (SD = 8.6 years). Table 1 summarizes this and other relevant descriptive information for the sample.
Table 1. Descriptive Characteristics of the FASCI Pilot Test Sample

<table>
<thead>
<tr>
<th></th>
<th>LA ((n = 43))</th>
<th>Noyce ((n = 2))</th>
<th>K-12 ((n = 8))</th>
<th>Faculty ((n = 3))</th>
<th>Grad Students ((n = 9))</th>
<th>Entire Sample ((N = 65))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Age (SD)</strong></td>
<td>20.33 (2.50)</td>
<td>22.5 (0.71)</td>
<td>46.25 (5.75)</td>
<td>57.0 (11.53)</td>
<td>24.44 (3.25)</td>
<td>25.85 (11.48)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>53%</td>
<td>47%</td>
<td>50%</td>
<td>50%</td>
<td>0%</td>
<td>33%</td>
</tr>
<tr>
<td><strong>Mean Years</strong></td>
<td>1.22 (2.33)</td>
<td>5.00 (5.66)</td>
<td>15.75 (6.30)</td>
<td>30.33 (17.50)</td>
<td>1.67 (1.58)</td>
<td>4.53 (8.59)</td>
</tr>
<tr>
<td>Teaching Experience (SD)</td>
<td>Astronomy 9%</td>
<td>33%</td>
<td>8%</td>
<td>Biology 12%</td>
<td>8%</td>
<td>Chemistry 12%</td>
</tr>
</tbody>
</table>

We purposefully chose our sample in order to gather a heterogeneous set of respondents. We hypothesized that very novice teachers (such as LAs) would have a limited repertoire of instructional strategies than more experienced teachers such as practicing K-12 teachers or Faculty Experts. Similarly, we expected novices such as graduate students in the STEM disciplines to be less student-centered in their instructional approaches as compared to more experienced K-12 teachers and Faculty Experts.

**Development of the FASCI Instrument**

The foundation for the FASCI development has been Wilson’s (2005) construct
modeling approach to measurement\textsuperscript{3}. In this approach, one begins by establishing a developmental theory for the latent construct to be measured. The instantiation of this theory is known as a construct map. Next, a set of items is designed to elicit information about how much or how little of the construct any given respondent possesses. A scoring rule is subsequently established for the anticipated item responses. After collecting data through pilot tests or field tests, patterns of scored responses across examinees and items are modeled statistically to make inferences about the location of a respondent (and item) on the construct map. The precision of the inference is evaluated by its associated estimate of measurement error. Each step of the construct modeling approach typically leads to iterative revisions for the construct map, item design, scoring rules, and measurement model. Indeed the results we present below obscure the many iterations and revisions that preceded them. In what follows we present the current state of our thinking about the FASCI instrument and its score interpretations.

The construct map in Figures 1 and 2 posits two distinguishing features of teachers’ strategic knowledge as they engage (or do not engage) in the formative assessment process. First, formative assessment depends on a teacher’s attention to student ideas. This constitutes the student-centered instruction (SCI) dimension of a teacher’s strategic knowledge (Figure 2). Second, the formative assessment process depends on a teacher’s attention to student learning outcomes and use of such information to modify their strategies when these strategies do not appear to be working. This constitutes the flexible application (FA) dimension of a teacher’s strategic knowledge (Figure 1). We would expect that the two dimensions in the construct map of strategic knowledge to be positively correlated since a teacher who is high on the SCI dimension is by definition one that is attentive to student conceptions and ideas, and therefore more likely to use such information to modify their instructional strategies than a teacher with a more behavioral view of teaching. As is evident in Figures 1 and 2, each dimension of the construct map is characterized by a rather small number of qualitative levels. This is not meant to suggest that more levels do not exist, only that to this point we were unable to establish additional indicators that could be used to distinguish them reliably.

\textsuperscript{3} This approach can also be mapped onto the “Assessment Triangle” presented in *Knowing What Students Know* (Pellegrino, Chudowsky, Glaser, & National Research Council Division of Behavioral and Social Sciences and Education Committee on the Foundations of Assessment, 2001). We start with a theory of respondent cognition (in this case, the learning progression), proceed to gather some observations of that cognition (through responses to the instrument), and use a measurement model to interpret those responses.
<table>
<thead>
<tr>
<th>Level</th>
<th>Respondent Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td></td>
</tr>
</tbody>
</table>
| 2     | • The teacher has repertoire of strategies that can be used to facilitate student learning within a given class session.  
      | • If the teaching strategy comprised of these acts is not producing the desired result, sometimes it can be modified.  
      | • The teacher recognizes that the choice of a class activity and associated teaching strategy will depend upon variables specific to the classroom context. |
| 1     | • The teacher has a repertoire of strategies that can be used to facilitate student learning within a given class session.  
      | • If an activity based on a particular teaching strategy is not producing the desired result, the activity can be modified by selecting a different strategy. |
| 0     | • The teacher has a limited repertoire of strategies.  
      | • Once a particular activity has been selected for a class session, it is not easily modified with a different strategy. |

*Figure 1. Flexible Application Dimension of Strategic Knowledge*
Level | Respondent Characteristics
---|---
1 | • Discussion of interactive teaching which would be *observable* to the teacher or to an outside “other.”
   Teacher | Students
   and/or
   Students | Students

0 | • No discussion of interactive teaching
   • Teacher primarily views classroom activity as a way to help students make sense of new ideas. Information goes from teacher to student.
   Teacher | Students

*Figure 2. Student-Centered Instruction Dimension of Strategic Knowledge*

The FASCI construct map shown in Figures 1 and 2 represent our attempts to reliably measure (a) the type of instructional strategy invoked by a teacher for a given classroom session, and (b) the extent to which this strategy is modified after student learning outcomes have been assessed (whether the assessment is formal or informal). Below we discuss how we are defining and distinguishing instructional strategies, what it means for an instructional strategy to be modified, and what it means for an instructional strategy to be classified as “student-centered.”

We define a “strategy” as a teaching act chosen purposefully to engage students with the content or skills they are expected to learn. So while all strategies are acts of instruction, not all acts necessarily constitute instructional strategies as we are defining the term. For example, the use of humorous anecdotes by a teacher can go a long way towards establishing a classroom environment that is conducive to learning, but we would not characterize this as an instructional strategy in and of itself. Treagust (2007) distinguishes six categories of instructional strategies commonly used in science teaching:

1. Demonstrations
2. Explanations (e.g., telling students the answer, interactive explanations)
3. Questioning (e.g., dialog, debate, discussion, use of wait-time)
4. Representations (e.g., concrete, verbal, mathematical, visual, symbolic, gestural)
5. Group and Cooperative Learning (e.g., jigsaw, peer-tutoring, group investigations)
6. Inquiry (e.g., exploration, conceptual invention, conceptual expansion)
All teachers have a potential repertoire consisting of these different instructional strategies. While any given classroom activity will usually involve some combination of strategies, one will usually be most prominent. Treagust (2007) presents the categories above as existing on a continuum from “teacher-centered” to “student-centered,” with demonstrations being the least student-centered and inquiry being the most student-centered. For FASCI responses, we say that an instructional strategy has been modified when the respondent articulates a shift in emphasis within a given item response from one instructional strategy to another (i.e., from a demonstration to questioning). The key criteria we use for determining whether a strategy is student-centered is whether the strategy is being used to implement a classroom activity that enables students to be actively engaged with the concepts at hand, and that provides an opportunity for student to express their developing ideas and conceptions.

The FASCI construct map represents an initial hypothesis about the factors that distinguish teachers that are more sophisticated in their strategic knowledge about teaching from those that are less sophisticated. The next step in developing an operational instrument was to create items capable of eliciting information about what teachers do (and the reasons why) when they are in the formative assessment cycle.

**Designing Scenario-based Items**

A construct map is a first step in operationalizing hypothesized characteristics of teachers with different levels of strategic knowledge. The items on a survey instrument are the means by which information about a given respondent’s location on the construct map is elicited. In the present assessment context the optimal choice of items to pose—both in terms of content and format—is not at all straightforward. Our approach has been to develop standardized items that present teachers with a variety of classroom-based scenarios. For each scenario, we provide three prompts that allow teachers to elaborate on their instructional decisions. The general structure for each item is shown below.

<table>
<thead>
<tr>
<th>Step 1: Present a classroom scenario that involves a specified learning activity. The scenarios should vary to the extent that they appear more student-centered or teacher-centered.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item Prompt (a) asks: How might this activity facilitate student learning?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2: Present information hypothetically elicited from the learning activity posed in the scenario that represents a potential obstacle to student learning.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item Prompt (b) asks: Describe both what would you do and what you would expect to happen as a result.</td>
</tr>
</tbody>
</table>

| Item Prompt (c) asks: If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session? |
The item below illustrates how the template above is applied:

<table>
<thead>
<tr>
<th>Students are working in groups of four to discuss a conceptual question you provided them at the beginning of class.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a)</strong> How might this activity facilitate student learning?</td>
</tr>
<tr>
<td>As the activity proceeds, one group gets frustrated and approaches you—they’ve come up with two solutions but can’t agree on which one is correct. You see that one solution is right, while the other is not.</td>
</tr>
<tr>
<td><strong>b)</strong> Describe both what would you do and what you would expect to happen as a result.</td>
</tr>
<tr>
<td><strong>c)</strong> If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?</td>
</tr>
</tbody>
</table>

While we have developed a total of 15 different scenarios, the pilot testing of the FASCI that we discuss below contains only the following five:

Scenario 1) Students are working in groups of four to discuss a conceptual question you provided them at the beginning of class.

Scenario 2) You are working out an example problem up on the board.

Scenario 3) You have just finished giving a lecture on a complicated topic.

Scenario 4) You have given your students a quiz to assess their understanding of a difficult topic.

Scenario 5) In talking with one of your students you discover that they have a misconception about a central topic presented in that week’s class. You attempt to address the misconception by having a one-on-one conversation with the student.

The following contextual constraint is given at the beginning of each scenario: “Please assume (unless it is otherwise specified) that you are teaching a high school course in physics, chemistry, biology, earth science or math to a class of 25-30 students.” The use of these classroom scenarios is intended to make it harder for respondents to provide what they consider to be the desired answer. In addition to these scenario-based items, a set of background questions are included about a respondent’s academic background and previous teaching experiences. Many of the scenarios used in the FASCI instrument were inspired by the actual experiences recounted to us by Learning Assistants through interviews. All scenarios were written by a team of faculty, graduate students, and K-12 teachers with expertise in both measurement and science education. The full version of the FASCI used in pilot testing is shown in Appendix A.

---

4 Although we recognize the potentially problematic nature of using the term “misconception” rather than “alternative conception” or “prior ideas,” we chose to use it in this particular item because it reaches a larger audience of respondents when interpreted in its colloquial meaning.

5 We have found that it takes the average respondent about 30 minutes to complete the most recent version of the FASCI.
**Scoring Rules for FASCI Items**

We developed two separate rubrics with rules for scoring every set of three item responses to each scenario on the FA and SCI dimensions. These scoring rules (shown in Appendix B) are based on the construct maps in Figures 1 and 2. SCI scores are based solely upon responses to prompt (a), while FA scores are based on the combination of responses to prompts (b) and (c). After an initial classroom scenario has been presented, prompt (a) asks respondents how the activity described in the problem statement (which features an implicit instructional strategy) might facilitate student learning. A response is given an SCI score of “1” rather than “0” only if there is some explicit evidence that the teacher views the potential for learning in this activity from a student-centered standpoint. In such responses, the classroom activity is viewed as an opportunity for the teacher to interact with the students, for the students to interact with one another, or for individual students to express or defend their ideas.

After prompt (a), new information is provided in prompt (b) about a potential obstacle to learning under the initial activity. After describing this obstacle, prompt (b) asks the respondent what he or she would do and what he or she would expect to happen as a result. This marks the first opportunity for a teacher to choose an instructional strategy beyond the one provided in the initial scenario. We do not however, consider this response a modified strategy, but rather a baseline choice of strategy. It is in prompt (c)—when the respondent is told that the approach he or she described in (b) has not produced the expected result—that we look for evidence that a teacher can modify his or her baseline instructional strategy. If the response indicates a modified instructional strategy (e.g., shift from questioning to collaborative learning as described in Treagust’s (2007) categories) in prompt (c), the FA score moves from a “0” to “1.” If the respondent then **qualifies** the modification with one or more conditional statements (typically based on contextual features), the FA score moves from a “1” to a “2.” In short, there are three score levels on the FA dimension. At the lowest level, there is no indication that a teacher is flexible in their application of instructional strategies; at the middle level, the teacher indicates some flexibility; at the highest level, the teacher explicitly indicates both flexibility, and an awareness of the context for being flexible in their application of different strategies. At the highest score level we are trying to capture the type of teacher that recognizes that there will always be a context when a given instructional strategy, no matter how carefully considered, may fail. Such teachers, we suspect, will tend to answer prompt (c) with “it depends” and then explain why.

In summary then, each scenario on the FASCI instrument results in a separate item score for each dimension. The pilot test version of the FASCI contained 5 scenarios; hence there were 5 items with a maximum raw score of 5 points for the SCI dimension, and 5 items with a maximum raw score of 10 points for the FA dimension. These raw scores were subsequently modeled using item response theory (IRT), which we describe below.

**Pilot Testing**

In our recent pilot test of the FASCI, the survey was administered online (via [www.questionpro.com](http://www.questionpro.com)) to the sample described above. The online administration included a respondent consent form. Raw data was downloaded from the web interface, cleaned, and
imported into a Microsoft Access database for storage and scoring. The database interface uses a set of forms within which raters score each set of responses.

Two part-time graduate students (mentioned above) were trained to assist in scoring. Along with the other graduate student researcher, all pilot test responses were scored by these three raters. Disagreements between raters were moderated and consensus scores were used in subsequent qualitative and quantitative analyses.

**Papers, Talks, and Workshops**

During the 2007-2008 academic year, the research team engaged in three separate dissemination activities:

1. Workshop given March 1 at the 2008 conference of the Physics Teacher Education Coalition (PTEC), Austin, TX (Talbot et al., 2008). The purpose of this workshop was to introduce the FASCI instrument to PTEC member institutions and recruit participants for future administrations of the survey. We also solicited and received feedback on the design of the scenario-based items and scoring of the responses.

2. Paper presented March 28 at the 2008 Annual Meeting of the American Educational Research Association (AERA) (Talbot & Briggs, 2008). The subject of this paper was the preliminary analysis and results from the recent FASCI pilot test.

3. Manuscript submitted to the *Journal of Research in Science Teaching* (JRST). The subject of this paper was a more comprehensive description and analysis of the FASCI pilot test, as well as an attempt to conceptualize the FASCI as a learning progression for teacher strategic knowledge (Briggs et al., 2008).

**Findings**

**Qualitative Analysis**

We initially made detailed qualitative observations of FASCI item responses to convince ourselves in a holistic sense (before applying our specific scoring rules) that a range of sophistication exists in responses to the FASCI with respect to the factors we envisioned on our construct map. Below are examples of responses which we consider to be examples of responses on the SCI and FA categories respectively.

**Example of High SCI:**

Scenario 2: You are working out an example problem up on the board.
Prompt a): How might this activity facilitate student learning?

High SCI Response: “They are able to see your own meta cognitive processes, allowing them to think about a problem step by step instead of just a problem and an answer. They can also deter [sic] from your model and speak up and say how they see the problem and how they do it, which might help out other students that think about it differently than you do as well.” (ID = 2001949)

This respondent appears to envision the students as actively engaged in the learning process. Even when the scenario provided involves a scenario that might not appear amenable to student-centered instruction (the teacher doing an example problem on the board), the respondent discusses student involvement and how the students might help each other to understand the topic under study.
Example of High FA:
Scenario 1: Students are working in groups of four to discuss a conceptual question you provided them at the beginning of class.
Prompt c): If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?
High FA Response: “Depends on the question and its role in the class. If the question was fundamental and something we needed to build on right away, it's conceivable that I might *discuss* the two solutions, have OTHER groups help argue out which was right (and why) and which was wrong (and why). If time was short, I might just present such a 'discussion' myself. If there was time and other groups were also struggling on the same point, I might build a secondary activity to approach the same concept from a new angle.” (ID = 2191029)

This respondent mentions the context dependence (questions’ role in the class, time available) as determining in part what instructional strategy they will take next. It also appears that the respondent has a large repertoire of teaching activities from which to draw, thereby increasing his/her potential to be flexible.

The following are example responses in which SCI and FA were rated as low.
Example of Low SCI:
Scenario 2: You are working out an example problem up on the board.
Prompt a): How might this activity facilitate student learning?
Low SCI Response: “I would do this so that students would have a record in their notes of how to solve the problem the correct way. Also if they see the problem done on the board and they copy it down, that is two chances for them to remember the correct way to solve the problem.” (ID = 1961454)

This response is from an individual who views the activity as an opportunity for students to gain the knowledge from an expert. The respondent does not envision students as being actively engaged in the presentation of this example problem, in contrast to the high SCI example presented above.

Example of Low FA (score = 0):
Scenario 4: You have given your students a quiz to assess their understanding of a difficult topic.
Prompt c): If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?
Low FA (score = 0) Response: “I would give a brief but thorough explanation of the topic at hand, and then offer a 'make-up' quiz. Then I would repeat the previous process, if necessary.” (ID = 2003336)

This respondent does not modify their teaching approach based on prompt (c) – they try the same thing over again and do not discuss any contextual factors which may bear upon their decision to do so.
Quantitative Analysis of FASCI Item Responses

Following preliminary qualitative analyses, responses to the FASCI items were independently rated by three trained raters using detailed scoring guides and supporting documents that had evolved from earlier pilot tests and their associated scoring moderation sessions. The percent agreement between raters on FA scores across five tasks ranged from 61% to 91% with an average of about 79%. The percent agreement on SCI scores across five tasks ranged from 45% to 95% with an average of about 67%. The variability in the score agreement across tasks suggests that considerable work will need to be done to improve the training of raters, although this problem is mitigated to the extent that all raters score the same respondents and all discrepant scores are resolved. This was the approach taken for the scoring reported here. The frequency distribution of final item scores by dimension for all 65 respondents is provided in Table 2. Note that there is missing data for some dimensions on some items. Approximately 5% of the surveys were not entirely completed.

Table 2. Frequency of Scores for each item on each Dimension (N = 65)

<table>
<thead>
<tr>
<th>Item</th>
<th>FA Dimension Score</th>
<th>SCI Dimension Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>15%</td>
<td>75%</td>
</tr>
<tr>
<td>2</td>
<td>41%</td>
<td>54%</td>
</tr>
<tr>
<td>3</td>
<td>25%</td>
<td>68%</td>
</tr>
<tr>
<td>4</td>
<td>40%</td>
<td>53%</td>
</tr>
<tr>
<td>5</td>
<td>36%</td>
<td>55%</td>
</tr>
</tbody>
</table>

We modeled the two dimensions of the FASCI separately using the Partial Credit Model (Masters, 1982), and estimated associated item and person parameters using the software ConQuest (Wu, Adams, & Wilson, 1997). In general, the match between the responses predicted by the model and the item scores observed was quite good. The Wright Maps for each dimension are displayed in Figures 3 and 4. The score scale on the Wright Map is represented in logit values that range from negative to positive; these values mark the location of both respondents and items on a common scale. Lower or higher values represent respondents with less or more sophistication, and items with less or more difficulty. The raw score associated with each logit value is provided along the right side of the continuum. Mean group locations are denoted by group labels. A histogram distribution of respondents is shown to the right of the logit values, with the location of any single respondent represented by each “X”. The mean person sophistication estimate for the FA dimension is just below -1 logit, which corresponds to a raw score of 3.8 points (out of 10 possible); for the SCI dimension, mean person ability is about 0 logits, which corresponds to a raw score of 2.7 points (out of 5 possible).

---

6 The fit of the category characteristic curves for each item was evaluated using Weighted Mean Square Error fit statistics (Bond & Fox, 2001; Wilson, 2005; Wright & Masters, 1982)
<table>
<thead>
<tr>
<th>Logit(raw)</th>
<th>Group Mean</th>
<th>Respondents</th>
<th>Step 1</th>
<th>Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>(10)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9)</td>
<td>Faculty Experts</td>
<td>item 2</td>
<td>item 3</td>
</tr>
<tr>
<td>2</td>
<td>(8)</td>
<td>X</td>
<td>item 1</td>
<td>item 5</td>
</tr>
<tr>
<td>1</td>
<td>(7)</td>
<td>Noyce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>(6)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5)</td>
<td>K-12, Grad Students</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>(4)</td>
<td>LA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>item 2 item 4</td>
<td>X</td>
<td>item 5</td>
</tr>
<tr>
<td>-2</td>
<td>(2)</td>
<td>X</td>
<td>item 3</td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>(1)</td>
<td>X</td>
<td>item 1</td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>(0)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each X represents 1 respondent, each row is 0.25 logits

Item 1: Students working in groups and come up with 2 solutions- they can’t agree on which is correct
Item 2: Teacher working a problem on the board and makes a mistake
Item 3: Teacher gives a presentation on a complicated topic and students look confused
Item 4: Teacher gives a quiz and the students do poorly
Item 5: Talking one-on-one with a student who has a misconception

Figure 3. FA Wright Map
Figure 4. SCI Wright Map

Each X represents 1 respondent, each row is 0.255 logits

- Item 1: Students working in groups and come up with two solutions - they can’t agree on which is correct
- Item 2: Teacher working a problem on the board and makes a mistake
- Item 3: Teacher gives a presentation on a complicated topic and students look confused
- Item 4: Teacher gives a quiz and the students do poorly
- Item 5: Talking one-on-one with a student who has a misconception
The location of estimated item difficulty parameters is indicated on the right side of each map. For the FA dimension, there are two difficulty thresholds represented for each item. The level 1 or 2 threshold marks the point on the logit continuum where there is a 50% probability for a respondent to receive a score above the lowest FA category. The level 2 threshold marks the point on the logit continuum where there is a 50% probability for a respondent to receive a score in the highest FA category (level 2). For the SCI dimension, there is one difficulty threshold represented for each item because it is scored as 0 or 1. According to the model, a respondent at the threshold location for a given item on the Wright Map has a 50% probability of receiving a score in category 0 or category 1 on that particular item.

The interpretation of these Wright Maps depends on the relative location of respondents and item difficulty thresholds. The higher the location (i.e., estimated sophistication) of a respondent relative to the location (i.e., estimated difficulty) of the items, the more we are likely to infer that a given respondent has a high level of strategic knowledge on the given dimension. So, for example, consider the two respondents represented by shaded X’s in Figure 3. One of these shaded respondents in Figure 3 has an FA ability estimate of 2 logits, and according to the model this person has a 53% chance of receiving the highest FA score on items 1 and 5, a 43% chance of receiving the highest FA score on item 4, a 40% chance of receiving the highest score on item 3, and a 33% chance of receiving the highest score on item 2. In contrast, the shaded respondent with an FA ability estimate of -2 logits has only a 2% chance of receiving the highest FA score on item 5, but a 46% chance of receiving a score of 1 on that item.

One noteworthy finding that the FA Wright Map helps make visually apparent is the large gap between the difficulty thresholds for each item: This is indicative of the small number of respondents (see Table 2) who give a response to prompt c that takes context into account when discussing their modification of an instructional strategy. Hence for most respondents, it was relatively easy to give a response scored in the middle FA category (a “1”) relative to the lowest (a “0”), but it was considerably harder to give a response that was scored in the highest FA category (a “2”). An examination of the SCI Wright Map reveals that two out of five items (items 1 and 5) were relatively easy for respondents. The easiest item in which respondents demonstrate student-centered thinking is item 1, in which all but three respondents received a score of 1. This is not entirely surprising, as we would expect the scenario of “students working in groups” to be an easy situation for an individual to conceive of as student-centered or as an opportunity for interactive teaching. The next easiest was item 5, in which the scenario states that “you are discussing a misconception with a student in a one-on-one conversation.” Again, we would expect most respondents to conceive of this situation as an opportunity for interactive teaching. The most difficult SCI item is the scenario in which the respondent is told that they have “just finished giving a lecture on a complicated topic” (item 3). Again, we would expect this scenario which presents a formal, teacher-led presentation would be a more difficult one to conceptualize as a student-centered instructional activity. Overall, the SCI items span a range of difficulty which includes the ability estimates of all respondents in the pilot sample.

The locations of respondents on the FA and SCI continua represented in the Wright Maps are estimates, and as such they contain some degree of measurement error. This measurement error can be quantified for each respondent by what is known as the standard
error of measurement (SEM). If scores from the FASCI are to be used for summative purposes, then it is important that the SEM is relatively small. Across the range of estimated respondent locations on both the FA and SCI dimensions, the SEM is about 1 logit. To evaluate the magnitude of the SEM, we can contrast it to the range of observed respondent locations. For the FA dimension this range is about 7 logits (from -4 logits to 3 logits); for the SCI dimension this range is much less, about 4.2 logits (from -2 logit to 2.2 logits). This implies that classification of respondents into statistically distinguishable categories is currently more plausible on the basis of FA scores than it is for SCI scores. Using a more formal approach described by Wright & Masters (1982, p. 105-106) one could conclude that there are at least 3 statistically distinct groupings (or person “strata”) of respondents on the basis of estimated locations on the FA dimension, but only 1 on the SCI dimension. A traditional way to summarize the proportion of score variability that is “real” and not due to measurement error is with a reliability coefficient. The reliability coefficients for the FA and SCI dimensions are 0.71 and 0.47 respectively.

To what extent are the distributions of respondent sophistication estimated on the FA and SCI Wright Maps consistent with the discrete levels that were hypothesized in the FA and SCI dimensions of the FASCI construct map? The answer for the SCI dimension is relatively straightforward at this point: FASCI items are not yet capable of reliably distinguishing respondents on this dimension—almost half of the observed variability on SCI items can be attributed to measurement error. Because of this, there is little point in trying to map SCI scores back to the SCI dimension of the FASCI construct map. On the other hand, at least from a statistical perspective, it is possible to reliably distinguish three levels on the FA Wright Map. This judgment might be analogous to the standard setting activities conducted to establish cut-points for the scores on state-level standardized tests considered to represent, for example, “proficient” performance. If such an approach were to be taken, one would need to pick two cut-points on the FA Wright Map that distinguish respondents whose collective pattern of item responses was reflective of levels 0, 1 or 2 on the FA dimension. For the sake of illustration, the FA cut-point differentiating a 0 from 1 might be set at -2 logits, and a 1 from a 2 might be set at 2 logits. Another possibility, if the underlying sample of teachers had been sampled such that they were representative of some well-defined population, would be to establish the cut-points on the basis of norm-referenced criteria.

**Evidence for Validity of Score Interpretations**

There are two hypotheses about FASCI score interpretations that we can begin to evaluate with our pilot test data. First, we have assumed that the scores for each dimension would be positively correlated. A scatterplot of FA and SCI scores is provided in Figure 5, and provides evidence for a positive linear relationship with a correlation coefficient 0.59, a value that actually understates the magnitude of the relationship because it is attenuated by measurement error.

---

7 In IRT models, the SEM actually varies as a function of ability, so the SEM of 1 reflects the typical value found. The range of observed SEM values is between 1.05 and .84 for the FA dimension, and between 1.2 and .90 for the SCI dimension.
Second, we have assumed that higher scores on the FASCI dimensions are indicative of substantive distinctions among teachers. Recall that there were 5 subgroups to consider:

1. Learning Assistants ($n = 43$)
2. Noyce Fellow ($n = 2$)
3. Practicing K-12 Teachers ($n = 8$)
4. Faculty Experts ($n = 3$)
5. Graduate Students ($n = 9$)

The location of each of these groups is shown on the FA and SCI Wright Maps. While these results are hardly conclusive given the small and limited sample we have gathered, they are suggestive nonetheless. Because of the observed unreliability of SCI scores, we focus on group comparisons for the FA dimensions. We note that the relatively inexperienced LAs (with a mean teaching experience of 1.2 years) score lower on the FA dimension than the more experienced Noyce Fellows (mean teaching experience of 5 years) or the Faculty Experts (mean teaching experience of 30 years). Because Noyce Fellows take on increasing levels of responsibility, and because they are enrolled in education courses in which they have many hours per week working with K-12 students, we would expect them to score higher on the FA scale than the LAs. One surprising finding is that our sample of practicing K-12 teachers (mean teaching experience of 15.8 years) also scored rather low on the FA dimension. In fact, they are located at essentially the same place as the Graduate Student teaching assistants, who we expected to be low due both to their lack of experience (mean of 1.7 years) and their lack of teacher education. We
had hypothesized that the FA scores of K-12 teachers would be higher based in part on their extensive classroom experience. One possible explanation for their low scores could have to do with their teacher education and/or induction programs; another possible explanation is that our interpretation of these scores in invalid. The best way to evaluate this would be to observe the same teachers in their classrooms, a validation activity we are planning to conduct in the future.

**Future Directions**

The research questions established for the CTL team in the LA-TEST grant were as follows:

1. What is the effect of the LA treatment on the sophistication of pedagogical understanding?
2. Does sophistication of pedagogical understanding vary by length of exposure to the LA treatment?
3. How is the sophistication of STEM LAs qualitatively distinct from the sophistication of STEM students who did not participate as LAs?

To address these questions it has been necessary to develop and begin validation activities around a survey instrument capable of measuring “sophistication of pedagogical understanding.” This instrument, which had the acronym SPUnC in the original proposal, has been renamed the FASCI (see above). Considerable evidence as to the validity of FASCI score interpretations has been established during the first two years of the project, and we will continue to gather more evidence along these lines in year 3. We plan to investigate the degree to which the content-neutrality of the current FASCI scenarios may pose a threat to the validity of score interpretations. The current FASCI scenarios were designed in this way so that we could administer the instrument to teachers from across the science disciplines. However, we are mindful of the fact that strategic knowledge for science teaching is often thought of as being specific to the type of science being taught (e.g., physics vs. biology). We are currently developing a new version of the FASCI in which the scenarios are contextualized within the content of physics. Specifically, we will investigate the following research questions:

1. How do scores on the existing content-neutral FASCI compare to scores on the physics-specific FASCI for both physics experts and novices?
2. To what extent is each version of the FASCI able to capture change in a respondent’s level of sophistication over the course of a semester?
3. Does the physics-specific FASCI provide a more reliable measure of teacher strategic knowledge than the content-neutral FASCI? Also as mentioned above, a validation activity that we will undertake in the coming year involves addressing the following question:
4. How do actual teaching practices compare to individual’s scores on both the physics-specific and content-neutral versions of the FASCI?

To investigate this latter question, we will work with the LATEST K-12 research team and observe former LAs and Noyce fellows in their initial teaching assignments. We plan to
collect interview data and scores from observation protocols (specifically, the Reformed Teaching Observation Protocol (RTOP; Sawada, Piburn, Judson, Turley, Falconer, Benford, & Bloom, 2002)) and compare that data to scores on both the existing content-neutral and new physics-specific versions of the FASCI for these same individuals.

As of year 3 of this project we are also now prepared to begin addressing our three principal research questions with FASCI scores as the outcome variables. We will be gathering a sample of respondents as described in the LA-TEST proposal and also implementing the analytical approach presented in sections 4.2.2 and 4.2.3. The will be no changes to the staffing of the project in the coming year.
References


Thank you for agreeing to respond to the FASCI. Your participation should take about 45-60 minutes of your time.

The first set of questions pertains to your personal information

2) Which statement best describes you?
   a) I am a Learning Assistant
   b) I am a Noyce Fellow
   c) I am a student in the teacher education program, but not in the LA program
   d) I am STEM Faculty
   e) I am School of Education Faculty
   f) I am a STEM graduate student
   g) I am a School of Education graduate student
   h) I am a Practicing K-12 teacher

3) What is your age?

4) What is your gender?
   a) Female
   b) Male

5) What is your race?
   a) American Indian or Alaska Native
   b) Asian
   c) Black or African American
   d) Hispanic or Latino
   e) Native Hawaiian or Other Pacific Islander
   f) White

The next set of questions pertains to your most recent teaching experience

6) What was the setting of your most recent teaching experience
   a) Elementary School (K-5)
   b) Middle School (6-8)
   c) High School (9-12)
   d) Instructor for an undergraduate course
   e) Learning Assistant for an undergraduate course
   f) Teaching Assistant for an undergraduate course
   g) Instructor for a graduate course
   h) Teaching Assistant for a graduate course
   i) Other (Please specify): ___________________________
7) In what content area were you teaching?
   a) Astronomy
   b) Biology or Life Science
   c) Chemistry
   d) Engineering
   e) Geology or Earth Science
   f) Physics
   g) Mathematics
   h) Statistics
   i) Other (Please specify)

8) What was the approximate number of students in your class(es)?
   a) Less than 10
   b) 11-20
   c) 21-30
   d) 31-50
   e) 51+

9) What was the gender distribution of your class(es)?
   a) More females than males
   b) About equal distribution of females and males
   c) More males than females

10) What was the racial/ethnic composition of your class(es)?
    a) Primarily African-American
    b) Primarily Asian
    c) Primarily Hispanic
    d) Primarily white/Caucasian
    e) Other (Please specify):

11) How many years of experience have you had in the role of a teacher, in any setting?
    Please specify: _____
For the scenarios that follow, please assume (unless it is otherwise specified) that you are teaching a high school course in physics, chemistry, biology, earth science or math to a class of 25-30 students.

12) Students are working in groups of four to discuss a conceptual question you provided them at the beginning of class.

   a) How might this activity facilitate student learning?

   As the activity proceeds, one group gets frustrated and approaches you—they’ve come up with two solutions but can’t agree on which one is correct. You see that one solution is right, while the other is not.

   b) Describe both what would you do and what you would expect to happen as a result.

   c) If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?

13) You are working out an example problem up on the board.

   a) How might this activity facilitate student learning?

   You accidentally make a mistake in solving the problem but don’t realize this until you get to the end of your solution and realize that the answer doesn’t make sense. No one in the class has said anything, so you’re not sure if they caught the mistake or not.

   b) Describe both what would you do and what you would expect to happen as a result.

   c) If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?

14) You have just finished giving a lecture on a complicated topic.

   a) How might this activity facilitate student learning?

   You notice that many of the students in the class have very confused expressions on their faces.

   b) Describe both what would you do and what you would expect to happen as a result.

   c) If the approach you described above in (b) didn’t produce the result(s) you anticipated by the end of that class session, what would you do in the next class session?
15) You have given your students a quiz to assess their understanding of a difficult topic.

   a) How might this activity facilitate student learning?

       Many of your students are discouraged after performing poorly on the quiz.

   b) Describe both what would you do and what you would expect to happen as a result.

   c) If the approach you described above in (b) didn’t produce the result(s) you
       anticipated by the end of that class session, what would you do in the next class
       session?

16) In talking with one of your students you discover that they have a misconception about
a central topic presented in that week’s class. You attempt to address the misconception
by having a one-on-one conversation with the student.

   a) How might this activity facilitate student learning?

       Despite your conversation, the student maintains the same misconception.

   b) Describe both what would you do and what you would expect to happen as a result.

   c) If the approach you described above in (b) didn’t produce the result(s) you
       anticipated by the end of that class session, what would you do in the next class
       session?
Appendix B

FASCI Scoring Instructions and Detailed Scoring Guides

Scoring Instructions

Every item response (consisting of responses to prompts a, b, and c) gets 2 scores:

SCI (Student-Centered Instruction) score

After reading the responses to all prompts, you will use the SCI scoring guide to assign a score for the SCI dimension

FA (Flexible Application) score

After reading the responses to prompts b and c, you will use the FA scoring guide to assign a score for the FA dimension
Scoring Guides

Flexible Application (FA)

<table>
<thead>
<tr>
<th>Score (level on construct)</th>
<th>Modification of teaching approach</th>
<th>Discussion of contextual factors that bear on the modification of the teaching approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>1</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>0</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Criteria for FA scoring:
- Modification of teaching approach (drawing from their repertoire)
  - Note: the modification needs to take the activity beyond the original activity presented in the scenario. They need to “break out” of the loop.
- Discussion of contextual factors that bear on the modification of the teaching approach

*Note: for FA, it doesn’t matter if their teaching strategies are student-centered or teacher-centered, just that they modify or change their approach somehow.

Student-Centered Instruction (SCI)

<table>
<thead>
<tr>
<th>Score (level on construct)</th>
<th>Discussion of interactive teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>YES</td>
</tr>
<tr>
<td>0</td>
<td>NO</td>
</tr>
</tbody>
</table>

Criteria/notes for SCI scoring:
- Discussion of interactive teaching in the learning activity. The teacher views the students’ role in the activity from a constructivist standpoint. They see the learning activity as an opportunity for interacting with the student so that they can identify the students’ current conceptions, understanding of the material, or level of engagement with the material. This may include students interacting with the teacher, articulating and defending their ideas with others, asking questions, working on a problem, manipulating apparatus, etc. The teacher takes action to involve the student in the learning process, such as questioning them in order to identify their previous/current conceptions.
Discipline-Based Educational Research (DBER) Team

Summary

Benjamin Spike

Activities

The DBER team consists of five STEM departments (Physics, APS, Applied Math, MCDB, and Chemistry) that have agreed to study the effects of educational reform on their lower-division courses. All five of the DBER departments have introduced undergraduate Learning Assistants (LAs) in recent years as a course reform. Other individually-instituted reforms include the use of tutorials, clickers, LA-led review sessions, and oral examinations.

The DBER team meets on a weekly basis during the school year, and includes members from all participating STEM departments as well as the School of Education. A typical meeting includes about 25 faculty members. The DBER team remains active in the research community through talks, publications, and workshops, which are documented in their respective individual reports.

Findings

The primary research questions for the DBER team concern the performance of students and student educators in LA-supported science courses:

3. **How do LAs compare to other STEM majors in terms of their content understanding, beliefs about the discipline, and beliefs about learning in the discipline?**

   Since LA selection is based partially on performance in lower-division courses, they tend to enter with comparatively high conceptual and belief survey scores. Nonetheless, the DBER team demonstrates continued learning gains and positive belief shifts as a result of the LA experience. Departments that examine LA conceptual knowledge report that some LAs have even been observed to out-perform graduate TAs on conceptual surveys. Longitudinal studies have also shown that students who were LAs tend to perform better than non-LAs in upper-division coursework.

   The DBER team has also demonstrated increased LA interest in education, through survey data and in the form of participation in volunteer-based K-12 outreach programs. Some have even chosen to return to the LA program on a volunteer basis. This observation has, in part, inspired the creation of the CU Teach program, in order to encourage and recruit STEM students with an interest in science education.

   The primary tool for measuring educational and disciplinary beliefs is the CLASS survey, which began as a physics survey and has been adapted to suit various courses. The
use of conceptual surveys is less uniform, however; some departments continue to use established conceptual surveys, while others have not developed one. Standardizing conceptual evaluation across the DBER disciplines remains a challenge due to the great diversity of subject matter.

4. **What effects can be observed on student achievement in courses that are supported by LAs?**

The DBER team continues to observe measurable, positive effects on students in reformed courses during AY 07-08. These effects include improved learning gains, decreased failure rates, and greater student satisfaction with coursework.

In some cases, the recent adoption of LAs affords no measurable change compared to previous semesters, and tools are being developed to more carefully observe the effects. Negative shifts in beliefs have also been observed to carry over from traditional to reformed courses, which remains an active subject of inquiry.

Examples of effects observed across the DBER team for AY 07-08:

- In physics, normalized learning gains continue to be higher than in traditionally taught courses. The effect appears to be dependent on the instructor’s familiarity with PER methods and reforms.
- Significant drop in failure rate during AY 07-08 as a result of reforms in Calculus I & II. In Spring 08 the failure rate of Calc II fell to 17% for the first time, from 31% and 27% in previous years.
- MCDB students who experienced weekly interactions with LAs had significantly higher learning gains and post test scores, and performed better in the class overall than their peers.
- More positive student course rating in APS, with specific mention of help provided by LAs.

Departments have continued to expand the number of LA-supported courses, and the number of professors requesting LA support increases with each academic year. In addition, all departments report increased departmental support in the form of funding, building space, and instructors. The favorable results observed in lower-division courses have also prompted departments to examine teaching methods in upper-division courses, as was done with PHYS 3310 in AY 07-08. The findings strongly suggest that the effects of these reforms are not limited to introductory coursework.

Individual findings from each discipline are included in the reports that follow.
Discipline-Based Educational Research (DBER) Team

Physics

Steve Pollock and Noah Finkelstein

Activities

Department Info

Transformed course details:

**PHYS 1110 (Physics I, Mechanics)**
Enrollment typically 500-600 each semester. There are three 50 minute lectures a week (using Peer Instruction) and one 50-minute Tutorial section for ~24 students, led by a grad physics TA and 1-2 LAs.
Principal role of LA is to facilitate UW-style Tutorials, (four per week), along with office hours in the Physics help room.

**PHYS 1120 (Physics II, E&M)**
Enrollment typically 350-500 each semester. There are three 50 minute lectures a week (using Peer Instruction) and one 50-minute Tutorial section for ~24 students, led by a grad physics TA and 1-2 LAs
Principal role of LA is to facilitate UW-style Tutorials, (four per week), along with office hours in the Physics help room.

**PHYS 1010/1020 (Physics of Everyday Life)**
Enrollment typically 200 each fall for 1010, and 50 each spring for 1020. There are three 50 minute lectures a week (using Peer Instruction). Principal role of LA is to facilitate student conversations during lecture periods, along with running optional study/homework sessions

**PHYS 2130/2170 (Modern Physics)**
Enrollment 70-200 in 2130 (aimed at Engineers), 40-80 in 2170 (aimed at Physics majors) There are three 50 minute lectures a week (using Peer Instruction)
Principal role of LA is to facilitate student conversations during lecture periods, along with running optional study/homework sessions

**PHYS 4810/7810 (Teaching and Learning Physics)**
Enrollment typically 12-24, mixed graduate and advanced undergraduate. Meets twice a week for 75 minutes. (Not run in the 2007-2008 academic year, scheduled to run Fall 2008) Enrolled students work in roles parallel to those of LAs.
PHYS 3310  (Junior level Electricity and Magnetism)
As part of a project to transform an upper-division physics course building on principles of education research, a former physics LA worked with a lead faculty member, developing and implementing Tutorial activities.

Current course syllabi can be found at
http://www.colorado/edu/physics/phys[1110, or 1120, etc]
For earlier semesters, the naming convention is e.g.
http://www.colorado/edu/physics/phys1110/phys1110_fa04

The Physics department is a member of the Physics Teachers Education Coalition (PhysTEC, http://www.ptec.org/)

Faculty for AY 2007-2008 in courses involving LAs:

Fall 2007:
Phys 1010 (Everyday life): Margaret Murnane. 1 LA (returning)
Phys 1110 (Mechanics, calc-based): Charles Rogers (lead) and Heather Lewandowski (in charge of LAs). 7 LAs, of whom 1 was returning
Phys 1120 (E&M I, calc-based): Steven Pollock (lead) and Victor Gurarie (in charge of LAs). 7 LAs, of whom 3 were returning
Phys 2130 (modern for engineers): Dan Dessau. 1 LA, returning
Phys 2170 (modern for physics majors) Noah Finkelstein. 2 LAs

Spring 2008:
Phys 1020 (Everyday life): Noah Finkelstein. 2 LAs
Phys 1110 (Mechanics, calc-based): Kevin Stenson (lead) and Kathy Dessau (in charge of LAs). 7 LAs, of whom 1 was returning
Phys 1120 (E&M I, calc-based): Heather Lewandowski (lead) and Charles Rogers (in charge of LAs). 7 LAs, of whom 3 were returning
Phys 2130 (modern for engineers): KT Mahanthappa. 1 LA, returning
Phys 3310 (Electricity and Magnetism): Steven Pollock. 1 former LA (working for independent study credit)

Applications for Fall 2007: 42 new and 6 returning (of the 48 total, 18 had another departmental interest as well) Females: 13, plus 4 returning.

Applications for Spring 2008: 52 new and 9 returning (of the 61 total, 32 had another departmental interest as well) Females: 15, plus 2 returning.

Accepted: Fall 2007: 18 LAs (6 female) (6 returning, 2M, 2F)
Accepted: Spring 2008 17 LAs (5 female) (6 returning, 4M, 2F)
Longer term brief summary of LAs in Physics:
We have hired a total of 132 undergraduate LAs since Fall 2003.
Of these, 34 were female = 26%
34 of these total LAs have served two or more semesters. (14 of these were female)

The rough total number of applicants for physics LA positions over 10 semesters is 413. Roughly 80 of these also applied to another departments too.
Of the 413 total, approximately 94 were female. (23%)
Findings

LAs

Content expertise

Physics I: FMCE (Force and Motion Concept Evaluation) a research-based, published assessment tool, given pre and post in Phys 1110 (introductory, calculus based physics):

Conceptual Learning: Learning Assistant data (not all matched - typically missing ~1 or 2 for pre and/or post)

<table>
<thead>
<tr>
<th>Semester</th>
<th>N</th>
<th>Pre ave</th>
<th>Post ave</th>
<th>Gain of ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp06</td>
<td>5</td>
<td>98.3</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>Fa06</td>
<td>7</td>
<td>88.7</td>
<td>98.5</td>
<td>0.87</td>
</tr>
<tr>
<td>Sp07</td>
<td>7</td>
<td>87.4</td>
<td>94.8</td>
<td>0.59</td>
</tr>
<tr>
<td>Fa07</td>
<td>7</td>
<td>97.0</td>
<td>98.9</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Notes: We did not take FMCE data for LAs in earlier semesters. Given the consistently very high pretests, and essentially "pegged" posttests, for all the previous semesters, we have decided not to continue measuring LA learning in this environment with the FMCE instrument.

Physics II: BEMA(Brief Electricity and Magnetism Survey)

<table>
<thead>
<tr>
<th>Semester</th>
<th>N</th>
<th>Pre ave</th>
<th>Post ave</th>
<th>Gain of ave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fa04</td>
<td>8</td>
<td>.54</td>
<td>.73</td>
<td>.42</td>
</tr>
<tr>
<td>Sp05</td>
<td>6</td>
<td>.75</td>
<td>.90</td>
<td>.61</td>
</tr>
<tr>
<td>Fa05</td>
<td>8</td>
<td>.66</td>
<td>.76</td>
<td>.31</td>
</tr>
<tr>
<td>Sp06</td>
<td>5</td>
<td>.73</td>
<td>.81</td>
<td>.31</td>
</tr>
<tr>
<td>Fa06</td>
<td>6</td>
<td>.64</td>
<td>.71</td>
<td>.19</td>
</tr>
<tr>
<td>Sp07</td>
<td>4</td>
<td>.71</td>
<td>.80</td>
<td>.31</td>
</tr>
<tr>
<td>Fa07 (CSEM)</td>
<td>6</td>
<td>.73</td>
<td>.87</td>
<td>.51</td>
</tr>
<tr>
<td>Sp08</td>
<td>6</td>
<td>.70</td>
<td>.76</td>
<td>.20</td>
</tr>
</tbody>
</table>

LA posttest results fluctuate considerably, and with the small numbers of LAs, it is not clear what "error bar" to associate with these numbers.

Other courses:

Phys 1020 (Test constructed by C. Wieman and K. Perkins)
Spring 2008: Two learning assistants, pre average 20.5/26, post average 24/26, Normalized gain 63.6%.

There is presently no established pre/post exam for the 1010 course.
(We are still working on compiling LA pre/post data for the Phys 2130/2170 reformed courses)
Beliefs

CLASS data is not available for the LAs.

Other LA activities

1) LAs in physics assist with Tutorials and other transformed classes, but also they all work 1-2 hours/week in the Physics Help room. This is essentially "office hours" for any students in any of the lower division introductory courses, in a large space in the basement of Duane Physics

2) The JILA Molecular, Atomic, and Optical Physics NSF Physics Frontier Center outreach activities include recruiting and training university students to participate in inquiry-based science activities in K12 environments. We have started to use volunteer LAs as participants (alongside grad student volunteers) for the first time this (2007-2008) academic year. The JILA program recruited about 8 LA’s for various activities over the last year. These students tended to come back each semester, which could be considered one of the best indications that they find this non-paid, non-credit experience worthwhile in addition to their busy schedule. We are actively considering adding the JILA program to the list of possible required activities for returning LA’s beginning in Fall 2008.

LA-supported courses

General Introduction: Course development and transformations have been going on since before the LA-TEST grant, with our first Tutorials implemented in 2003, and data collection beginning in 2004. We have observed and documented the following:
- improve student content mastery in:
  - mechanics (Physics 1), electricity and magnetism (Physics 2), physics of everyday life, modern physics, and (when offered) teaching and learning physics (Phys 4810/7810)
- often succeed at supporting favorable attitudes and beliefs:
  - there are some instances of degradation on the CLASS (phys 1), similar to other published work (See Adams, Phys Rev 2, 010101)
  - several of the implementations have seen no degradation on the CLASS, which we consider significant success
- provide lasting impact on content mastery
  - in upper division physics we observe difference between those students who have had Tutorials and those who have not
  - there is differential positive impact for those students who were LAs
  - outcomes apparently depend upon which / how faculty implement these curricula
Content Conceptual Learning: Data for classes as whole:

Physics I: (Phys 1110) (introductory, calculus based physics)

FMCE (Force and Motion Concept Evaluation) given pre and post:

<table>
<thead>
<tr>
<th>Semester</th>
<th>Recitation</th>
<th>N (matched)</th>
<th>Ave pre*</th>
<th>Ave post*</th>
<th>Norm. gain**</th>
</tr>
</thead>
<tbody>
<tr>
<td>F01</td>
<td>Traditional</td>
<td>265</td>
<td>36</td>
<td>52</td>
<td>0.25</td>
</tr>
<tr>
<td>F03 (PER)</td>
<td>Tutorials</td>
<td>400</td>
<td>53 (FCI data)</td>
<td>81 (FCI data)</td>
<td>0.63</td>
</tr>
<tr>
<td>S04 (PER)</td>
<td>Tutorials</td>
<td>335</td>
<td>28</td>
<td>74</td>
<td>0.64</td>
</tr>
<tr>
<td>F04</td>
<td>Workbooks</td>
<td>302</td>
<td>33</td>
<td>69</td>
<td>0.54</td>
</tr>
<tr>
<td>S05</td>
<td>Traditional</td>
<td>213</td>
<td>28</td>
<td>58.5</td>
<td>0.42</td>
</tr>
<tr>
<td>F05</td>
<td>Traditional</td>
<td>293</td>
<td>32</td>
<td>58</td>
<td>0.39</td>
</tr>
<tr>
<td>S06</td>
<td>Tutorials</td>
<td>278</td>
<td>26.5</td>
<td>59.5</td>
<td>0.45</td>
</tr>
<tr>
<td>F06</td>
<td>Tutorials</td>
<td>331</td>
<td>33</td>
<td>67</td>
<td>0.51</td>
</tr>
<tr>
<td>S07</td>
<td>Tutorials</td>
<td>363</td>
<td>29</td>
<td>62</td>
<td>0.46</td>
</tr>
<tr>
<td>F07</td>
<td>Tutorials</td>
<td>336</td>
<td>32</td>
<td>69</td>
<td>0.54</td>
</tr>
<tr>
<td>S08</td>
<td>Tutorials</td>
<td>434</td>
<td>29</td>
<td>51.5</td>
<td>0.32</td>
</tr>
</tbody>
</table>

* Average(%) is scored using recommended FMCE rubric, for students with matched pre-post data.

(Standard error of the mean is between 1-2% for all terms). F03 used the FCI exam pre and post, all other terms are FMCE. (Note similar gains for F03 and S04 on these two exams.)

** Normalized gain $\langle g \rangle$ in the last column is computed as the gain of the average pre and post scores for matched students. (Standard error of average gains is roughly ±.02 for all terms.)

FMCE pre-post results: Fa08 was taught by experienced faculty, Sp08 by new faculty. These outcomes represent nearly historical extremes for FMCE results at CU (with LAs), but both show learning gains well above the expected 0.2-.25 observed in Hake's metastudy for traditional classes.
Physics II: (Phys 1120) (introductory, calc-based second semester physics)

BEMA(Brief Electricity and Magnetism Survey) pre and post:

### Summary of BEMA scores

<table>
<thead>
<tr>
<th>Semester</th>
<th>N (matched)</th>
<th>posttest score (%)</th>
<th>Normalized gain &lt;g&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>F04 (PER)</td>
<td>319</td>
<td>59</td>
<td>0.44</td>
</tr>
<tr>
<td>S05 (PER)</td>
<td>232</td>
<td>59</td>
<td>0.43</td>
</tr>
<tr>
<td>F05</td>
<td>325</td>
<td>50</td>
<td>0.33</td>
</tr>
<tr>
<td>S06</td>
<td>196</td>
<td>53</td>
<td>0.37</td>
</tr>
<tr>
<td>F06</td>
<td>351</td>
<td>56</td>
<td>0.40</td>
</tr>
<tr>
<td>S07</td>
<td>261</td>
<td>53</td>
<td>0.37 *</td>
</tr>
<tr>
<td>F07 (PER)</td>
<td>161 (BEMA)</td>
<td>61</td>
<td>0.47 **</td>
</tr>
<tr>
<td>F07 (PER)</td>
<td>170 (CSEM)</td>
<td>66</td>
<td>0.50</td>
</tr>
<tr>
<td>S08</td>
<td>316</td>
<td>58</td>
<td>0.41</td>
</tr>
</tbody>
</table>

All semesters shown used Tutorials in recitation. Standard errors of the mean are 1-2% on posttest, and roughly ±.02 for normalized gains. (PER) in parentheses means the lead faculty that term was a member of the Physics Education Research group. (Pretests are very steady at 26±1%)

* No pretest given in S07, we used the but ave pretest (26±1%) to estimate normalized gain.

** In F07, half the class took BEMA, half CSEM  (For the CSEM, the pretest score was 32% )

Histogram of BEMA results for most recent two semesters. Both semesters involved experienced faculty, both used similar curriculum and materials. The reproducibility is striking. The BEMA is harder than the FMCE, only scattered national results are published, but these outcomes look very high compared to e.g. original BEMA (and CSEM) publications.
Phys 2130/2170 (Modern physics, for engineers/physics majors respectively)

We have administered the QMCS (Quantum Mechanics Concept Survey) pre and post for several semesters in both of these classes. The table shows historical results for both courses, in some cases (reformed, using Learning Assistants) and Trad (more traditional instruction, without LAs) versions of these courses.

<table>
<thead>
<tr>
<th>Course: Phys 2130 (engineers)</th>
<th>Style</th>
<th>Pre</th>
<th>Post</th>
<th>&lt;g&gt;</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp05</td>
<td>Trad</td>
<td>NA</td>
<td>51</td>
<td>NA</td>
<td>68</td>
</tr>
<tr>
<td>Fa05</td>
<td>Reformed</td>
<td>32</td>
<td>69</td>
<td>.54</td>
<td>162</td>
</tr>
<tr>
<td>Sp06</td>
<td>Reformed</td>
<td>30</td>
<td>65</td>
<td>.49</td>
<td>156</td>
</tr>
<tr>
<td>Fa06</td>
<td>Reformed</td>
<td>34</td>
<td>63</td>
<td>.45</td>
<td>73</td>
</tr>
<tr>
<td>Sp07</td>
<td>Reformed</td>
<td>36</td>
<td>62</td>
<td>.41</td>
<td>120</td>
</tr>
<tr>
<td>Fa07</td>
<td>Trad</td>
<td>34</td>
<td>54</td>
<td>.31</td>
<td>40</td>
</tr>
<tr>
<td>Sp08</td>
<td>Trad</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Course: Phys 2170 (phys majors)</th>
<th>Style</th>
<th>Pre</th>
<th>Post</th>
<th>&lt;g&gt;</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp05</td>
<td>Trad</td>
<td>NA</td>
<td>63</td>
<td>NA</td>
<td>64</td>
</tr>
<tr>
<td>Fa05</td>
<td>Trad</td>
<td>40</td>
<td>52</td>
<td>.21</td>
<td>54</td>
</tr>
<tr>
<td>Sp06</td>
<td>Trad</td>
<td>44</td>
<td>64</td>
<td>.37</td>
<td>23</td>
</tr>
<tr>
<td>Fa06</td>
<td>Trad</td>
<td>38</td>
<td>52</td>
<td>.22</td>
<td>54</td>
</tr>
<tr>
<td>Sp07</td>
<td>Trad</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Fa07</td>
<td>Reformed</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Sp08</td>
<td>Trad</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Comments: Pre and post scores are %. The standard error of the mean on the gains is typically of order 5%.

Pre is about 7 points higher for the physics majors than the engineers.

Post Trad is on average 12 points lower than Post Reformed in 2130

Post Trad is on average 8 points lower than Post Reformed in 2170 (although this course has only been taught once in the reformed style)

The survey was not given in either class in Sp08 (both traditional)
**Phys 3310:** (upper division Electricity and Magnetism for majors)

Our Learning Assistant (Darrin Tarshis) developed ten optional weekly Tutorials, using research-based and other reform materials obtained from Bruce Patton (Ohio state) ("Jackson by Inquiry", which he is modifying into "Griffiths by Inquiry"), Oregon State (OSU) Paradigms activities, and materials of Paul Van Campen (Dublin City University).

**Evaluations:**
1) A new upper-level conceptual post-test was developed for this course, the CUE (Colorado Upper-division Electrostatics assessment) Fourteen post-test questions were given in common in both semesters, and graded on a common rubric by two independent graders. See http://www.colorado.edu/sei/fac-resources/course-archives.htm
   
   **Traditional average:** 43±3.7    **Transformed average:** 61±4.0
   
   The difference is significant (p<0.001) and remains even if the portions of the exam requiring explanations (which were not stressed in the Traditional course, and that many students in that course left blank) are removed. Scores were normally distributed, and were low relative to exam scores because the test was scored more strictly than would an exam. Students in the Transformed course scored at least 10 points higher on each individual question on the test, save three (although only 8 of these differences were statistically significant due to low N) – a difference roughly equivalent to a letter grade.

2) Another comparison was made using more conventional exam questions:

![Plot shows common exam questions given in Fa07 (traditional) and Sp08 (reformed)](image)

Error bars indicate ± 1 SE. All differences are significant at p<0.05 level except Q5 and BEMA. The visible conclusion is that the transformed class (the course with clicker questions and LA-run Tutorials) generated higher results on all common conventional exam questions, not just on the conceptually focused questions of the CUE evaluation.
Other courses:

**Phys 1010 (Physics of Everyday Life)**
This course ran in Fall 2007 with one (returning) Learning Assistant (Christina Jones) who assisted in lectures and ran optional help/study sessions. There is no standardized pre-post test for this course presently in use. 81% of the students responded to an online end of term survey on their attitudes about the pedagogy used in the class. Students were asked about a variety of course elements (lectures, homework, clickers, help sessions) and asked to rate "how useful for your learning" and "how much did you enjoy". By far the most favorable responses were for the help/review sessions run by the LA. (<10% unfavorable, >50% favorable in both categories)

**Phys 1020 (Physics of Everyday Life)**
This course ran in Sp 2008 with two (returning) Learning Assistant who assisted in lectures and ran optional help/study sessions. A pre-post test developed by C. Wieman and K. Perkins was given: Matched (25 students out of 42) average pre is 10/26, post is 18/26 (Note: average posttest for all 42 students is 17/26) Normalized gain of these averages (matched only) is 49.6%

Older data from 2005 is available (also reformed course, with LAs), pre 9.8 -> post 15.8 Normalized gain was 37%

**Phys 4810/7810 (Teaching and Learning Physics)** was not offered in AY 2007-2008. (It will be offered next in Fall 2008, and is now established as a permanent physics department course satisfying upper division credit requirements)
### Attitudes and Beliefs: CLASS Data:

<table>
<thead>
<tr>
<th>Course/Term</th>
<th>N</th>
<th>CLASS overall positive</th>
<th>CLASS shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHYS 1110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phys 1110 Sp04 (PER)</td>
<td>398</td>
<td>68.3</td>
<td>+1.5 *</td>
</tr>
<tr>
<td>Phys 1110 Fa04 (No LAs)</td>
<td>389</td>
<td>65.8</td>
<td>-6.5 *</td>
</tr>
<tr>
<td>Phys 1110 Sp05 (No LAs)</td>
<td>349</td>
<td>62.7</td>
<td>-5.2 *</td>
</tr>
<tr>
<td>Phys 1110 Fa05 (No LAs)</td>
<td>277</td>
<td>69.5</td>
<td>-7.0 *</td>
</tr>
<tr>
<td>Phys 1110 Sp06</td>
<td>320</td>
<td>60.4</td>
<td>-3.3 *</td>
</tr>
<tr>
<td>Phys 1110 Fa06</td>
<td>360</td>
<td>68</td>
<td>-11.3 *</td>
</tr>
<tr>
<td>Phys 1110 Sp07</td>
<td>302</td>
<td>62.5</td>
<td>-3.6 *</td>
</tr>
<tr>
<td>Phys 1110 Fa07</td>
<td>346</td>
<td>65.3</td>
<td>-4.4 *</td>
</tr>
<tr>
<td>Phys 1110 Sp08</td>
<td>314</td>
<td>61.5</td>
<td>-6.2 *</td>
</tr>
<tr>
<td>PHYS 1120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phys 1120 Fa04 (PER)</td>
<td>326</td>
<td>70.0</td>
<td>-2.9 *</td>
</tr>
<tr>
<td>Phys 1120 Sp05 (PER)</td>
<td>219</td>
<td>66.4</td>
<td>-2.2 *</td>
</tr>
<tr>
<td>Phys 1120 Fa05</td>
<td>320</td>
<td>61.2</td>
<td>-3.9 *</td>
</tr>
<tr>
<td>Phys 1120 Sp06</td>
<td>193</td>
<td>63.3</td>
<td>-1.4</td>
</tr>
<tr>
<td>Phys 1120 Fa06</td>
<td>257</td>
<td>62.0</td>
<td>-0.8</td>
</tr>
<tr>
<td>Phys 1120 Sp07</td>
<td>225</td>
<td>63</td>
<td>-1.2</td>
</tr>
<tr>
<td>Phys 1120 Fa07 (PER)</td>
<td>182</td>
<td>62.7</td>
<td>+3.1 *</td>
</tr>
<tr>
<td>Phys 1120 Sp08</td>
<td>195</td>
<td>66.3</td>
<td>-0.9</td>
</tr>
<tr>
<td>PHYS 1020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phys 1020 Sp08 (PER)</td>
<td>28</td>
<td>62.2</td>
<td>-0.3</td>
</tr>
<tr>
<td>PHYS 2130</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phys 2130 Fa07</td>
<td>42</td>
<td>66.6</td>
<td>-3.2</td>
</tr>
<tr>
<td>PHYS 2170</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phys 2170 Fa07 (PER)</td>
<td>43</td>
<td>81.2</td>
<td>-6.0 *</td>
</tr>
<tr>
<td>Phys 2170 Sp08 (No LAs)</td>
<td>57</td>
<td>71.7</td>
<td>-5.0 *</td>
</tr>
</tbody>
</table>

Typical standard error of mean on these shifts is about +/-1 or 2%.

(PER) in parentheses means the lead faculty that term was a member of the Physics Education Research group.

* means that the shift is statistically significantly different from zero.

Note that many of these classes show overall negative shifts. These results are typical for attitude and belief shifts (as demonstrated e.g. by the MPEX)
Traditionally (as seen in the table on the previous page), Phys 1120 has shown small but statistically significant negative shifts on CLASS attitudes and beliefs. This year, we saw smaller than usual declines (and in Fall, slight increases in most categories, in some cases statistically significant) The course structure has not changed - it involves LAs in recitation/Tutorials, and clicker questions in lecture (although this year both terms had experienced and PER-aware faculty involved.) We do not know if these results are reproducible or to what we should attribute this change from historical trends.
**Evaluation of department buy-in/challenges**

**Buy-in**
- The physics department has typically been funding 1-2 LAs each term in addition to grant and/or institutional support used for the remaining LAs. The department has committed support for 2 LAs/term for the coming (08-09) academic year.
- The physics department partially supports a lead graduate TA whose primary additional responsibilities are to support faculty and LAs in the Tutorials.
- Starting Fall 2008, the Physics department has secured dedicated classroom space for the Tutorials in introductory physics. This involved a petition to the administration, which included evidence of learning gains in courses using Tutorials. This will allow us to release some of the temporary space in the basement labs we have been using for Tutorials.
- We have developed notebooks for new faculty in Phys 1110 and 1120 Tutorials, with materials, notes, and pretests, to make running this alternative pedagogy as smooth as possible.
- We have collections of concept tests (and other course materials, including online homework banks) which are provided to any interested new physics faculty when they are assigned to teach one of these transformed courses.
- We ran a faculty workshop in Fall 07 aimed at all new science faculty assigned to teach a course using Learning Assistants, to explain the program and help facilitate appropriate use of LAs.
- In Spring 2008, the physics course committee approved our petition to make "Teaching and Learning Physics" (an advanced undergrad/grad course on Physics Education Research and its applications) a regular physics course.
- The physics department became a member of the Science Education Initiative fall of 2007.
- The physics dep't hired a postdoc (Stephanie Chasteen) summer 2007 with SEI funds to work on Phys 3310 course transformations. She supervised an advanced LA in that course and is analyzing data on impacts.
- The physics department hired a second postdoc (Steven Goldhaber) summer 2008 with SEI funds to work on another upper division course.
- JILA and physics jointly hired a postdoc (Laurel Mayhew) who is working on outreach to local high schools.
- As of Spring 2008, over 20 different physics staff/faculty have been directly involved in teaching courses with LAs.

**Challenges**
- LA recruitment and hiring is a large chore which must be done every term. This has largely been done by two PER faculty members (SJP and MD), but it is time consuming with little personal reward, although of course it is essential to the running of the LA program in physics.
- Collection of pre/post data is another large chore which must be done every term. This again has been largely spearheaded by one PER faculty member (SJP), but requires cooperation of regular faculty, and is not easy to sustain or support. It is also quite easy to "slip" and miss collecting critical data, just by oversight.
References/sources for the various Posttests used:


**QMCS**: S. McKagan, see http://www.colorado.edu/physics/EducationIssues/QMCS/

**CUE**: S. Chasteen and S. Pollock, submitted to PERC 2008. See also http://www.colorado.edu/sei/fac-resources/course-archives.htm


Discipline-Based Educational Research (DBER) Team
Astrophysical & Planetary Sciences
Seth Hornstein and Doug Duncan

Activities

Department Info
In the academic year of 2007-2008, 5 faculty from the Department of Astrophysical and Planetary Sciences (Duncan, Stocke, Bagenal, Burns, & Hornstein) were involved with the LAs in 6 courses.

Transformed course details:

ASTR 1010 (Introductory Solar System Astronomy w/ Lab)
210 students in Fall 2007, 202 students in Spring 2008. Three 50 minute classes per week, plus 1 hr 50 min lab. LAs assist TAs with laboratory sections.

ASTR 1020 (Astronomy 2 (Stars & Galaxies) w/ recitation)
128 students in Fall 2007, 171 students in Spring 2008. Three 50 minute classes per week, plus 50 min recitation section. LAs lead weekly recitation sections of about 20 students. No tutorials; content is created by instructor and LAs. LAs also lead pre-exam review sessions.

ASTR 1110 (Introductory Solar System Astronomy)
211 total students in Spring 2008. Similar structure to ASTR 1010, but without recitations or laboratory. LAs assist with in-class tutorials and clicker questions. Also conduct end-of-term review sessions.

ASTR 2000 (Ancient Astronomies)
160 students in Fall 2007. Two 1 hr 15 min classes per week. LAs present researched material during class, in the form of PowerPoints or demonstrations.

Faculty for AY 2007-2008 in courses involving LAs:

Fall 2007:

ASTR 1010 (Introductory Solar System Astronomy w/ Lab): Fran Bagenal, instructor. ~5 LAs.
ASTR 1020 (Astronomy 2 (Stars & Galaxies) w/ recitation): Doug Duncan, instructor. 4 LAs.
ASTR 2000 (Ancient Astronomies): John Stocke, instructor. 2 LAs.
Spring 2008:
ASTR 1010 (Introductory Solar System Astronomy w/ Lab): Fran Bagenal, instructor. ~5 LAs.
ASTR 1020 (Astronomy 2 (Stars & Galaxies) w/ recitation): Jack Burns, instructor. 4 LAs
ASTR 1110 (Introductory Solar System Astronomy) (2 sections): Seth Hornstein, instructor. ~4 LAs.

Findings

LAs

Content expertise

1010: The LAs were used in 1010 primarily in the labs - as part of the learning team (instructor-TA-LA) for each lab section. The LA-TA meetings that I ran once a week suggested that the LAs helped bridge the cultural/academic gap between the students (non-scientists) and the TAs (mostly PhD candidates in APS). In Fall 2007, the instructor tried an experiment of the LAs running review sessions / study sessions in the dorms. This was a dismal failure. Very few (<20) students took up this opportunity. I think the reason is that by the time students realized they were struggling they decide either to punt the course (many going for course forgiveness, apparently) or had already gotten into a set lack-of-study-habit and were not willing/able to set up study sessions with the LAs.

2000: One LA learned how to use the planetarium projector to show basic motions in the sky. She gave review sessions in the planetarium and honed her skills in teaching modest sized classes. This LA is still somewhat shy and soft spoken in classes and so using her in this way was less intimidating than addressing the entire class. The second LA learned how to do library and internet research on topics related to the class and then made powerpoint presentations to the entire class. In this job she performed so well that she taught me some new material. Both LAs learned how to administer clicker questions and were involved in a few tutorials about good ways to pose clicker questions for maximum response. Both LAs met with students on an on-demand basis to review homeworks, prepare for examinations and to provide tutorial assistance for self-paced observing projects for the class. Both LAs attended lecture and critiqued the lecture with the instructor.

1020: LAs had oversight of 8 recitation sections (2 sections per LA). Although we met weekly to plan each week’s recitation content in general, the LAs had considerable freedom and authority to present the material according to her/his individual style. I saw dramatic improvement in the confidence of the LAs as the semester went forward. LAs become more creative, more interactive with students, and more successful in delivering on course material.
LAs were used in the lecture theater during peer discussion during clicker questions as well as during in-class tutorial usage. LAs sat in with groups, asked probing questions, and provided guidance when necessary. As the semester went on, LAs became more skilled at providing useful hints and stepping into groups already knowing what the difficulties were in the assignment. LAs met with the instructor and TA weekly to prepare for the upcoming tutorials, to discuss things seen in lecture and to share ideas. LAs also helped run (in pairs) final review sessions at the end of the semester.

Beliefs

All I have is anecdotal evidence that the LAs became more cautious about supporting poor behavior by their fellow students. They got to witness lack of attention in class, lack of preparation and studying for tests and general lack of enthusiasm first hand, as they will see in the future in their own classrooms.

LAs were much more comfortable approaching a group by the end of the semester. In the beginning of the semester, LAs would circulate aimlessly until flagged down by a group having trouble. By the end of the semester, LAs would join a group, unsolicited, and ask for students to explain their thinking.

LA-supported courses

Content

Instructor taught ASTR 1020 without LAs in Spring 2007 and with LAs in Spring 2008. Two sections always met at the same time, allowing LAs to support each other. This worked well. There was a dramatic improvement by a full 10 percentage points in each of the 3 midterm exams in 2008 versus 2007. The exams were comparable in content and question construction. Therefore, I believe this is evidence that the LAs and the recitation sections made a major improvement in student’s understanding and retention of course material.

The most successful and innovative thing the LAs did was in pre-exam review sessions. They broke their 20 students into groups of 3 or 4 and gave each a large sheet of paper ("butcher paper") and markers. Then they would state one of the key topics to be covered on the exam (e.g. "galaxy evolution"; "big bang") and have each group write three things they knew about the topic. The papers were then passed to another group whose job it was to evaluate what the previous group wrote – whether the statements were all correct. The LA acted as referee. This engaged far more students than a typical review with TA at the board and one question at a time being asked. In the future we will always do this kind of exam review.

I have implemented the self-paced observing projects since I began using LAs. The non-threatening presence of the LAs during observing project help
sessions has made these projects much more effective than they would be w/o LAs (although I have only my own personal perspectives to support this assertion...) I know that I would not have attempted the self-paced observing projects w/o LAs.

Beliefs

1020: Since this class topic is "Stars and Galaxies" the solar system diagnostic I began work on last summer, that Seth Hornstein is working on now, is not an appropriate instrument. However, I convinced Wieman Fellow Leilani Arthurs to do a major investigation with me on the topic of whether students become more "scientific" in their thinking after taking an introductory course. We gave all students the EBAPS instrument, and interviewed 40, split between my class and Seth Hornstein's ASTR1120 class as a control. His class does not have LAs but covers the same "Stars and Galaxies" topics.

The 9 interview questions were all concerned with the nature of science, doing science, and scientific thinking. Some were taken from the VNOS instrument. 40 interviews have all been transcribed this summer and Leilani and I are breaking them into categories and analyzing them. We have one major finding already: students in my class are far more likely than the control group to say that science is something anyone can do, you just have to work at it (as opposed to "science is something scientific people do."). I would expect that the emphasis on figuring things out for yourself that the LAs and I emphasize is a main reason for the difference. A second reason is that I explicitly discuss the nature of science and who does it and how during lecture.

Other impacts

1020: The FCQs for Spring 2008 were improved by a full point (from 4 to 5 on a scale of 6) from Spring 2007 in terms of the students' views of the class as a whole. Students specifically mentioned the impact of the recitations and the LAs in improving their understanding of difficult concepts. Students generally liked having LAs and small groups.

Students are divided nearly 50%-50% whether they like being "helped to figure out answers themselves" vs. just being told the answer. However, they understand (about 80% do) that they learn more this way. So here's an IMPORTANT RESULT: acceptance of the LAs by students is enhanced when the professor explains - many times, in class - why LAs are NOT giving students answers. When students understand that their learning and grades increase when they figure stuff out on their own they are happy. If they don't see the point they get mad from having to do work they expect their teacher or TA to do.

This is very similar to what we are finding in research on the use of clickers and peer instruction. Acceptance is much better when the faculty member clearly explains why students are being asked to do more of the hinking and explaining. Of course the exams better not be rote memorization if you tell students they need to learn to think and analyze.
Evaluation of department buy-in/challenges

- The ASTR 1020 course in the APS department has been redesigned to include a mandatory 1-hour small-group recitation section led entirely by LAs. Without LAs this small-group learning environment would be unsustainable.
- In order to foster a sense of community amongst the APS LAs, all LAs are given offices in the CASA stadium wing. These offices could be used for help sessions/office hours for the course they were acting as an LA as well as an on-campus home base for the LAs to call their own.
- The APS department has committed to funding 2 LAs with department funds for the 2008-2009 school year.
Discipline-Based Educational Research (DBER) Team

Applied Math

Mary Nelson, Anne Dougherty, Jim Curry

Activities

Department Info
The Department of Applied Mathematics has 15 tenured and tenure-track faculty, 5 instructors, and several lecturers who teach over 14,000 credit hours each year. In AY 2007-08 12 faculty used LAs in their classes. This is a significant increase from previous years and indicates the degree to which the LA program has become a part of the department.

Transformed course details:

**GEEN 1350** (Calculus I workshop)
Enrollment in the fall is 40-60 students and 20-30 in the spring. There is one class/week for 100 minutes. The workshop is led by one TA and 2 LAs. Students work in small groups on challenging problems to help them gain better conceptual understanding of the material. Principal role of the LAs is to facilitate small group discussions and help students think more deeply about the critical concepts. These LAs also help with homework and test grading (which helps them understand what students’ misconceptions are) and by the end of the semester they help with oral assessments.

**GEEN 1360** (Calculus II workshop)
Enrollment in the fall is 20-30 students and 40-60 in the spring. There is one class/week for 100 minutes. The workshop is led by one TA and 2 LAs. Students work in small groups on challenging problems to help them gain better conceptual understanding of the material. Principal role of the LAs is to facilitate small group discussions and help students think more deeply about the critical concepts. These LAs also help with homework and test grading (which helps them understand what students’ misconceptions are) and by the end of the semester they help with oral assessments.

**APPM 1340** (Calculus I, semester 1)
Enrollment in this class is about 100 students in three sections. There are 3 fifty minute classes a week. Learning assistants conduct 1.5 hour homework study sessions twice a week. Students are asked to put their questions on the board and help each other solve the problems. The principal role of the LAs is to ask probing questions to help the students face their misconceptions and to help them make mathematical connections. LAs also grade homework and exams and late in the first semester begin conducting oral assessments.
**APPM 1345 (Calculus I, semester 2)**
Enrollment in this class is about 90-95 students in three sections. There are 3 fifty minute classes a week. Learning assistants conduct 1.5 hour homework study sessions twice a week. Students are asked to put their questions on the board and help each other solve the problems. The principal role of the LAs is to ask probing questions to help the students face their misconceptions and to help them make mathematical connections. LAs also grade homework and exams and conduct oral assessments.

**APPM 3310 (Matrix Methods)**
Enrollment in this class is 70-80 in the fall (2 sections) and 60 in the spring semester (1 section). There are 3 fifty minute classes per week. Learning assistants conduct 1.5 hour homework study sessions twice a week. Students are asked to put their questions on the board and help each other solve the problems. The principal role of the LAs is to ask probing questions to help the students understand their misconceptions and to help them make mathematical connections. LAs also grade homework and assist with the exam grading.

**APPM 4350 (Fourier Series and Boundary Value Problems)**
Enrollment in this class was 76 in fall 2007 (2 sections). There are 3 fifty minute classes per week. Learning assistants conduct 1.5 hour homework study sessions twice a week. Students are asked to put their questions on the board and help each other solve the problems. The principal role of the LAs is to ask probing questions to help the students understand their misconceptions and to help them make mathematical connections. In addition, LAs provided a few office hours each week to work more intensively with individual students who were having difficulties. LAs also grade homework and assist with the exam grading.

**APPM 4520 (Mathematical Statistics)**
Enrollment in this class was 33 in fall 2007 and 33 in spring 2008. There are 3 fifty minute classes per week. Learning assistants conduct 1.5 hour homework study sessions twice a week. Students are asked to put their questions on the board and help each other solve the problems. The principal role of the LAs is to ask probing questions to help the students understand their misconceptions and to help them make mathematical connections. LAs also grade homework and assist with the exam grading.

Current course syllabi can be found on the Applied Mathematics home page:

http://amath.colorado.edu/courses/1340/
http://amath.colorado.edu/courses/1345/
http://amath.colorado.edu/courses/1350/
http://amath.colorado.edu/courses/1360/
http://amath.colorado.edu/courses/3310/
http://amath.colorado.edu/courses/4350/
http://amath.colorado.edu/courses/4520/
**Faculty** for AY 2007-2008 in courses involving LAs:

**Fall 2007:**

*APPM 1340:* (first of 2 semesters, Calculus I) Mary Nelson (lead) and Luis Melara, 3 LAs of whom all were new

*GEEN 1350 and GEEN 1360:* (Calculus I and II workshop) Mary Nelson (lead) and Adam Norris, 4 LAs of whom 3 were returning

*APPM 3310:* (Matrix Methods) Anne Dougherty (lead) and Greg Beylkin, 3 LAs of whom all were returning

*APPM 4350:* (Fourier Analysis) Mark Ablowitz, 1 LA who was returning

**Spring 2008:**

*APPM 1345:* (second of 2 semesters, Calculus I) Mary Nelson (lead), Anca Radelescu and Luis Melara, 4 LAs of whom 3 were returning

*GEEN 1350 and 1360:* (calculus I and II workshop) Mary Nelson (lead) and Adam Norris, 4 LAs of whom 2 were returning

*APPM 3310:* (Matrix Methods) Anne Dougherty, 2 LAs of whom both were returning

*APPM 4520:* (Mathematical Statistics) Tiejun Tong, 1 LA who was returning

**Applications** for Fall 2007: 27 new and 6 returning (of the 33, APPM was not the first choice of 7) Females: 6 new and 3 returning (about 27%); 10 were accepted, 3 females

**Applications** for Spring 2008: 26 new and 9 returning; (of the 37, 15 were only interested in APPM) Females: 9 new and 3 returning (about 35%); 11 were accepted, 4 females
Findings

LAs

- Content expertise

The Applied Mathematics department does not test the content knowledge of LAs. We do not have content tests available to us. Anecdotally, the lead faculty for each course are convinced that being a learning assistant brings about a substantial improvement in the LA’s content expertise. We rely on the LAs performance in Applied Math courses. We find that this is a good indication of a student’s content knowledge.

- Beliefs

We do not have CLASS data for the LAs.

- Other LA activities

The APPM LAs who work in GEEN 1350 and 1360 and the LAs who work in APPM 1340 and 1345 facilitate oral assessments every semester when they feel prepared to do so. They are required to help in their second semester as an LA. Students sign up for a voluntary, ungraded hour long session before each written unit test. The facilitator asks the students conceptual questions and students are helped to understand the important concepts underlying the material to be tested.

LA-supported courses

- Content

Success in APPM 1350 (Calculus I) using GEEN 1350 workshops:

Traditionally, the failure rate in APPM 1350, Calculus I, is 30-33%. For the first time in Fall 2007, the failure rate for Calculus I dropped to 22%. Our data supports the use of workgroups (GEEN 1350) with LAs.

The pie charts below show that the pass/fail rates for students attending workshops was 10% higher than that of non-workshop students.
Non-workshop pass/fail rates; 

Workshop pass/fail rates; 

LA-TEST Year 2 Report  
DBER: Applied Math  
78
The following chart shows that in fall 2007, students who participated in Orals for the first exam did significantly better in Calculus I. Those taking Orals averaged about B- for their overall grade. While those who did not take Orals averaged about a C.

<table>
<thead>
<tr>
<th>O1</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Oral 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam 1 (pre make-up)</td>
<td>333</td>
<td>17</td>
<td>100</td>
<td>75.09</td>
<td>15.006</td>
</tr>
<tr>
<td>Exam 2 (pre make-up)</td>
<td>328</td>
<td>17</td>
<td>100</td>
<td>75.04</td>
<td>15.112</td>
</tr>
<tr>
<td>Exam 3 (pre make-up)</td>
<td>324</td>
<td>15</td>
<td>99</td>
<td>65.31</td>
<td>18.609</td>
</tr>
<tr>
<td>Final Exam</td>
<td>330</td>
<td>0</td>
<td>149</td>
<td>88.07</td>
<td>30.568</td>
</tr>
<tr>
<td>finalgrade</td>
<td>324</td>
<td>.00</td>
<td>4.00</td>
<td>2.0352</td>
<td>1.28956</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>309</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Took Oral 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam 1 (pre make-up)</td>
<td>134</td>
<td>45</td>
<td>99</td>
<td>81.55</td>
<td>10.429</td>
</tr>
<tr>
<td>Exam 2 (pre make-up)</td>
<td>132</td>
<td>30</td>
<td>100</td>
<td>79.61</td>
<td>12.982</td>
</tr>
<tr>
<td>Exam 3 (pre make-up)</td>
<td>132</td>
<td>9</td>
<td>99</td>
<td>72.19</td>
<td>17.800</td>
</tr>
<tr>
<td>Final Exam</td>
<td>134</td>
<td>0</td>
<td>139</td>
<td>96.68</td>
<td>28.080</td>
</tr>
<tr>
<td>finalgrade</td>
<td>132</td>
<td>.00</td>
<td>4.00</td>
<td>2.6083</td>
<td>1.01358</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Success in APPM 1360 (Calculus II) using GEEN 1360 workshops:

The failure rate in APPM 1360, Calculus II, was 31% in Spring 2006. In Spring 2007, the failure rate dropped to 27%. For the first time in spring 2008, the failure rate for Calculus II dropped to 17%.

Our data supports the use of workgroups (GEEN 1360) with LAs. The pie charts below show that the pass rate for students attending workshops was 11% higher than that of non-workshop students.
The following chart shows that in spring 2008, students who participated in Orals for the second exam did significantly better on exams 1 and 2. The remaining data for this course is presently being analyzed.

<table>
<thead>
<tr>
<th>Orals Exam 2</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam 1 Grade</td>
<td>284</td>
<td>22</td>
<td>100</td>
<td>74.51</td>
<td>16.468</td>
</tr>
<tr>
<td>Exam 2 Grade</td>
<td>273</td>
<td>14</td>
<td>100</td>
<td>67.57</td>
<td>18.344</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam 1 Grade</td>
<td>132</td>
<td>30</td>
<td>100</td>
<td>80.15</td>
<td>13.617</td>
</tr>
<tr>
<td>Exam 2 Grade</td>
<td>129</td>
<td>19</td>
<td>100</td>
<td>75.22</td>
<td>15.216</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>129</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**APPM 3310, 4350 and 4520:**

No data has been collected yet for these courses since they have not used LAs long enough yet.

**Success in APPM 1340-45 (Calculus IA & IB) using Orals:**

Current data has not yet been analyzed, but previous semesters have demonstrated that students in the two-semester course pass over 90%. This is in contrast to similar at-risk students who pass at about a 10% rate in regular classes without Orals.

The following table shows how treatment students who were offered Orals did compared to control students who had no Orals. Treatment students were subdivided into those who are at-risk and those who were not at-risk. “At-risk” was determined by their placement scores at the beginning of the fall semester. If a student scored below 18 out of 30 on the placement test, they were considered at-risk. This test has been determined to be an excellent predictor of success rates in regular Calculus I classes. Control students were also subdivided into workshop and non-workshop students. No treatment students attended workshops.
The table below shows the effect size between control and treatment students.

<table>
<thead>
<tr>
<th>Comparison groups</th>
<th>Mean Exam Score</th>
<th>Standard deviation</th>
<th>Mean difference</th>
<th>Effect size in Stand. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment At-risk</td>
<td>93.13</td>
<td>27.414</td>
<td>39.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Workshop At-risk</td>
<td>53.93</td>
<td>20.949</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment At-risk</td>
<td>93.13</td>
<td>27.414</td>
<td>41.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Non-workshop At-risk</td>
<td>51.72</td>
<td>25.536</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Not-at-risk</td>
<td>98.20</td>
<td>25.302</td>
<td>24.22</td>
<td>.96</td>
</tr>
<tr>
<td>Workshop Not-at-risk</td>
<td>73.98</td>
<td>24.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Not-at-risk</td>
<td>98.20</td>
<td>25.302</td>
<td>20.03</td>
<td>.75</td>
</tr>
<tr>
<td>Non-wrkshp Not-at-risk</td>
<td>78.17</td>
<td>26.828</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment At-risk</td>
<td>93.13</td>
<td>27.414</td>
<td>14.96</td>
<td>.55</td>
</tr>
<tr>
<td>Non-wrkshp Not-at-risk</td>
<td>78.17</td>
<td>26.828</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Effect on students:**

Workshops based on the Treisman method, emphasis on correct placement of students and oral assessments have been the main transformations occurring in APPM courses, particularly the lower level courses.

- Improved pass rates in Calculus I and II. Traditionally, the failure rate for Calculus I in the Applied Mathematics department is about 30-33%. In the fall of 2007, the failure rate for Calculus I fell to 22%. In addition, the spring 2008 failure rate for Calculus II fell to 17%. The failure rate for Calculus II in Spring 2006 was 31% and in 2007 it was 27%.

**Evaluation of department buy-in/challenges**

**Buy-in**

- In the first years of the STEM grant only four faculty members participated.
  - those four faculty members continue to participate
    1. the chair, Jim Curry
    2. a senior instructor, Anne Dougherty
    3. two instructors; Nelson and Norris
  - presently, the following eight faculty have joined the effort
    1. four full professors: Harvey Segur, Mark Ablowitz, Greg Belykin and Congming Li
    2. one assistant professor: Tejun Tong
    3. three instructors: Anca Radelescu, Luis Melara, Ann Scheels
- numerous teaching assistants, learning assistants and Noyce fellows in the department have participated in the transformations

- Luis Melara, an instructor for one year, will take a tenure track position in the fall and has requested that we speak about Orals at a colloquium at his new college. He will offer Orals next year in his new position.
- We have created notebooks of APPM 1340 and 1345 Orals questions for the new faculty who will teach in subsequent semesters
- We will create notebooks of APPM 1350 and 1360 Orals questions for the new faculty who will teach in subsequent semesters
- We have developed notebooks of workshop questions for GEEN 1350 and 1360.
- We have developed 3 clicker concept questions for every Calculus I class period
- We have a book of clicker questions nearly ready for publication
- Based on the work we have done in transformation, we have applied for a CCLI Phase 2 NSF grant. We just submitted answers to questions that were raised by reviewers about the proposal.
- If our CCLI 2 grant is funded, oral assessments will be offered in Mechanical and Aerospace Engineering, and at a local high school. A Mechanical Engineering instructor has successfully implemented Orals into two of her classes already.
- In conjunction with the mathematics coordinator for the Boulder Valley School District, we will begin a two-year algebra course for struggling freshmen in fall 2008 that is based on our success with a two-semester Calculus I course. Our hope is that given two years, these students will master the material taught in a regular algebra course and then be able to take regular geometry and regular Algebra II courses instead of watered-down versions of those courses. This could allow these at-risk students the opportunity to pursue a college degree.
- The Applied Math department funded 3 LAs in fall 2007 and spring 2008. In addition, the department hired other undergraduates to work with the department’s learning assistants. These additional undergraduates did not take the LA course.
- In fall 2008, the department will hire at least 6 new learning assistants to be paid with course fees.
- In spring 2007, we allowed a former LA, Kate Spooner, to teach one Calculus I recitation in order to improve her teaching skills before student teaching.
- In fall 2007, we allowed one former LA, Amanda Geist, to teach two Calculus I recitations in order to improve her teaching skills before student teaching.
- In fall 2008, we intend to hire a former LA to teach 2 recitations.
- A former LA presented a poster about her LA experiences at the Joint Mathematics Meeting in January 2007.
- Two former LAs presented a poster about their LA experiences at the Joint Mathematics Meeting in January 2008.
- One of the full professors who taught an LA assisted course in spring 2008 hesitated to offer Orals. He did facilitate three Orals before the first test, and then sent an email to all his students before the second test saying that if they could do nothing else to prepare for the test, they should participate in an oral. He told them that a one hour oral would help them more than 3 hours of studying.
Applied Mathematics was invited to present information, a mock oral and data for the spring retreat for Presidential Teaching and Learning Scholars. The presentation was well received.

Three presentations were made at conferences about the transformations occurring in the department.

We were invited to speak about Orals at Boulder Valley School District. They plan to implement Orals in one high school in the district in Fall 2008. Three middle school teachers will use Orals as well.

We have been invited to submit an article to PRIMUS about our transformations. It should be submitted in the next week.

A full time instructor was hired to participate in the transformations. He taught the two-semester course, Calculus I and Calculus II. He participated in Orals in all of those courses.

One assistant professor who had an LA in the spring for the first time, has requested LAs for the coming year.

The College of Engineering supported our APPM 1340-45 classes in AY2007-2008, providing funds for a full time instructor.

The College of Engineering has provided rooms for the conduct of Orals throughout the year. They are offering additional rooms in fall 2008.

The Multicultural Engineering Program has repeatedly allowed us to use their conference room to conduct Orals. They have requested training for their homework study session facilitators.

**Challenges**

- Interviewing and hiring LAs each semester is a time consuming endeavor which we have not yet managed to streamline. It is largely accomplished by one instructor, which consumes a great deal of time for about two weeks every semester. It is necessary in order to hire LAs who will love their jobs and contribute positively to the APPM effort.
- APPM has not yet mastered use of the CLASS. We hope to give it pre and post in the fall semester.
- APPM does not have standardized pre/post tests for students or LAs. This challenge is being addressed this summer.
Discipline-Based Educational Research (DBER) Team

Molecular, Cellular, and Developmental Biology

Jenny Knight, Michelle Smith, Jia Shi

Activities

Department Info

Transformed course details:

**MCDB 1150** (Introduction to Cell and Molecular Biology)
374 students in Fall 2007. LAs were used to run voluntary recitation sections once per week. The sections had two different formats; peer-centered, LA-led working groups (56 students), and 1 TA-led study group (format more like a recitation section where TA presents material; 25 students).

**MCDB 1041** (Fundamentals of Human Genetics (non majors))
81 students in Fall 2007. On Fridays, LAs lead class in small groups (approximately 20 students per LA). Students worked in small groups to complete hands-on activities or worksheets prepared by instructor.

**MCDB 2150** (Genetics)
146 students in Fall 2007, 362 students in Spring 2008. LAs ran weekly group activities sessions during class. Once a week half the class would learn from group activities and the other half would learn from interactive lecture. The LAs facilitated student interaction in the group activities section and also reviewed the answers students handed in for the group activities.
LAs assisted with an experiment that examined how much students learned from peer instruction during clicker questions. Specifically, the LAs transcribed conversations around them and recorded how long peer instruction discussions were staying on task.
The LAs also ran optional homework study groups (6 different sections) that were regularly attended by approximately half of the class.

Faculty for AY 2007-2008 in courses involving LAs:

**Fall 2007:**
**MCDB 1150** (Intro to Cell & Molec. Bio.): Nancy Guild and Jennifer Martin, instructors; Jia Shi, STF. 4 LAs, of whom 3 were returning.
**MCDB 1041** (Mechanics, calc-based): Jennifer Knight, instructor. 4 LAs, all new.
**MCDB 2150** (Genetics): Ken Krauter, instructor; Michelle Smith, STF. 4 LAs, all new.
**Findings**

**LAs**

- **Content Expertise**

**Introduction to Molecular and Cellular Biology Assessment**

<table>
<thead>
<tr>
<th>Course</th>
<th>Number</th>
<th>Pre-assessment +/- SE</th>
<th>Post assessment +/- SE</th>
<th>Normalized Learning Gain +/- SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCDB 1150</td>
<td>4</td>
<td>67.7 +/- 2.9</td>
<td>80.6 +/- 1.0</td>
<td>37.3 +/- 9.3</td>
</tr>
</tbody>
</table>

**Genetics Assessment**

<table>
<thead>
<tr>
<th>Course</th>
<th>Number</th>
<th>Pre-assessment +/- SE</th>
<th>Post assessment +/- SE</th>
<th>Normalized Learning Gain +/- SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCDB 1041</td>
<td>4</td>
<td>72 +/- 11.4</td>
<td>86 +/- 7.6</td>
<td>50.8 +/- 11.9</td>
</tr>
<tr>
<td>MCDB 2150</td>
<td>6</td>
<td>78.7 +/- 6.8</td>
<td>88.7 +/- 4.9</td>
<td>54.2 +/- 13.3</td>
</tr>
</tbody>
</table>

(Note: the **MCDB2150 LAs outscored the MCDB2150 genetics TAs** (n= 11) at the beginning and the end of the semester: Pre-assessment score = 76.4 +/- 5.9; Post-assessment score = 88.0 +/- 3.4).

- **Beliefs**

We have been in the process of developing a belief survey for biology and do not have data on the LAs at this time. We did, however, administer the Bio CLASS to the students in these courses (see data below).

**LA-supported courses**

- **Content**

**MCDB 1150 students**

**Introduction to Molecular and Cellular Biology Assessment**

<table>
<thead>
<tr>
<th>Course</th>
<th>Number</th>
<th>Pre-assessment +/- SE</th>
<th>Post assessment +/- SE</th>
<th>Normalized Learning Gain +/- SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCDB 1150 (entire class)</td>
<td>307</td>
<td>42 +/- 0.8</td>
<td>71.8 +/- 0.8</td>
<td>51.1 +/- 1.3</td>
</tr>
<tr>
<td>Peer-centered, LA-led working group</td>
<td>56</td>
<td>42 +/- 1.8</td>
<td>75.5 +/- 1.7</td>
<td>58.0 +/- 2.6</td>
</tr>
<tr>
<td>Interested non-participants (control)</td>
<td>17</td>
<td>40.6 +/- 2.7</td>
<td>65.1 +/- 3.7</td>
<td>41.2 +/- 7.1</td>
</tr>
<tr>
<td>Interested, non-participant (no show)</td>
<td>28</td>
<td>45.3 +/- 2.8</td>
<td>67.4 +/- 4.0</td>
<td>40.4 +/- 5.5</td>
</tr>
</tbody>
</table>
Fall 2007 was the first time LAs were used in our large introductory biology course. We decided to study how useful additional student working groups were in improving student learning. Students voluntarily signed up for one of 4 additional group work sessions, led by a pair of LAs. These students worked on homework and other hands-on activities during this one hour/week session. The group that had weekly interactions with LAs, performed significantly better on the post test (p<0.05), and had higher learning gains. Students also performed better in the class overall if they were part of the LA-led working group than if they were part of the control group.

<table>
<thead>
<tr>
<th>Course</th>
<th>Number</th>
<th>Pre-assessment +/- SE</th>
<th>Post assessment +/- SE</th>
<th>Normalized Learning Gain +/- SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCDB 1041</td>
<td>61</td>
<td>33.6 +/- 1.52</td>
<td>69.7 +/- 1.85</td>
<td>53.5 +/- 3.0</td>
</tr>
<tr>
<td>Fall MCDB 2150</td>
<td>107</td>
<td>41.8 +/- 1.4</td>
<td>77.8 +/- 1.5</td>
<td>62.3 +/- 2.3</td>
</tr>
<tr>
<td>Spring MCDB 2150</td>
<td>314</td>
<td>37.7 +/- 0.8</td>
<td>75.9 +/- 0.8</td>
<td>61.7 +/- 1.3</td>
</tr>
</tbody>
</table>

The MCDB 1041 non majors genetics course was transformed several years ago into an interactive, clicker and activity-based class with some lecture, and has used LAs since 2004. We do not have data on student performance pre-LAs, because we had not yet developed a content assessment tool. The students in this class have high learning gains on the genetics assessment tool (especially considering the tool was designed for majors).

The fall MCDB2150 genetics majors course met three times per week. On Mondays and Wednesdays the students would attend lecture in a traditional lecture hall and answered approximately three clicker questions per class. On Fridays the students were randomly split into two equal-sized groups and the class met in a new interactive classroom with round tables. Group #1 participated in an interactive lecture that had approximately six clicker questions per class. Group #2 participated in group activities that were facilitated by LAs, TAs, and the course instructors. The content was the same for both group #1 and #2 and half way through the semester, the groups switched teaching approaches.

We found that students learn the same content knowledge in the interactive lecture and the group activities sections. In order to arrive at this conclusion, we measured learning at various time points during and after class, and only saw a significant difference during in-class clicker quizzes for the first half of the course. There was no significant difference between the two groups for all of the exam scores.

For the spring MCDB2150 course, we found that students learn from peer discussion with their neighbors during clicker questions and can apply that knowledge to a new question on the same concept. The LAs observed that
students had longer and more meaningful discussions on the medium and difficult questions when compared to the easy questions.

❖ **Attitudes and Beliefs**

We have been in the process of developing a biology-specific attitude assessment tool called the BioCLASS which is similar to the PhysicsCLASS. We have had over 80 experts respond to the statements and over 1,000 students from several CU biology courses take the assessment. We are currently using factor analysis to establish the categories (ex. connection to real world, problem solving ability). When we look at overall shifts for our intro courses, we see slight negative shifts (table below). We are working with the MCDB faculty to identify the statement or groups of statements that are causing the most concern and will develop a plan to address these statements in the upcoming school year.

**CLASS Data, Overall Shift:**

<table>
<thead>
<tr>
<th>Course/Term</th>
<th>N</th>
<th>CLASS overall agreement with experts</th>
<th>CLASS shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCDB 1150 Intro Fall 2007</td>
<td>215</td>
<td>70.8%</td>
<td>-9.2 ±1.5 (SE)</td>
</tr>
<tr>
<td>MCDB 2150 Genetics Fall 2007</td>
<td>104</td>
<td>76.6%</td>
<td>-4.2 ± 1.2 (SE)</td>
</tr>
<tr>
<td>MCDB 1041 Non majors Genetics Fall 2007</td>
<td>42</td>
<td>62.6%</td>
<td>-.09+/-.2.3 (SE)</td>
</tr>
</tbody>
</table>

The students in these classes show overall negative shifts in attitude, although the non-majors' shift is very small. These results are similar to attitude and belief shifts in Physics (as demonstrated e.g. by the MPEX) and the Physics faculty at CU.

Of the 170 students that regularly participated in the LA-lead homework study groups for Genetics 2150, **only 2 students** said that they were not satisfied with their decision to join a study group. In Intro 1150, when students were asked the question “Overall, how satisfied are you with the optional LA-led working group?” All students were either very satisfied (65%) or satisfied (35%).

Here are some of the positive things students had to say about their study/working groups:

- *We are able to all work together and actually understand each concept that is covered on the problem sets.*

- *I love the study groups. I learn the material much more thoroughly with my peers in an organized setting. Thank you for having them available to us!*
– It is impossible to do the homework without these study groups. We move through lecture material really fast and the study groups are a chance to slow down. This should be a required class....

Evaluation of department buy-in/challenges

Buy-in

- MCDB has been using LAs since before the inception of the LA-TEST program (first LAs used in 2003). Since then, the number of faculty using LAs has grown from 1 to 9. At least 4 other faculty have expressed an interest in using LAs in their courses in upcoming years. In addition to the LAs reported above, MCDB and the Science Education Initiative have funded LAs for two upper division courses in the department for two consecutive years: Immunology 4300 and Developmental Biology 4650. Because these are junior/senior level capstone courses, the LAs for these courses are typically seniors as well. The LAs in these courses help run optional study sessions, as well as being active participants in lecture (discussing clicker questions and guiding in-class activities).

- Because of the success of optional recitation sections this past year in both MCDB 1150 and 2150, the faculty voted to make such an experience available to all students. Beginning this fall, both MCDB 1150 and 2150 will have one-credit optional recitation sections that will be led by LAs. In these sections, students will work through homework assignments and group activities that are similar to the ones used this past year in both courses.

- We have archived all materials from the three courses described above, including clicker questions, homework questions, in-class activities, and pre/post assessments, so that others who may teach these courses in the future have access to all materials.

- Training the LAs to encourage peer-learning, and to successfully facilitate group work remains challenging. Despite the education seminar designed to train the LAs, we have found that they require considerable supervision, as well as attention to their content knowledge. We are working to develop new models of conducting LA meetings to help improve LA content knowledge and confidence.

- Several of our LAs have gone on to participate in education-related activities. Two LAs from this year’s pool are working as teachers for Teach for America next year. Another LA is working with her high school biology teachers to re-evaluate their curriculum. Two additional LAs will work on education-research related projects (funded by the SEI) this summer and into the next school year.
Discipline-Based Educational Research (DBER) Team

Chemistry

Laurie Langdon

Activities

Department Info

Transformed course details:

*CHEM 1111* (General Chemistry I)
Largest 5-credit course at CU. Each week, students attend three 50-minute lectures, one 50-minute recitation session, and one 3-hour laboratory session. (Typically, the 20-student recitation and lab occur in the same lab room, in a 4-hour block. Lab follows recitation.) Enrollment typically 800 – 900 students in Fall; 400 students in Spring. Two faculty members team-teach the three lecture sections in Fall; one faculty member teaches one lecture section in Spring. Graduate TAs lead lab sections, and co-lead recitation sections with LAs.

<table>
<thead>
<tr>
<th>Semester</th>
<th>LA role(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fa06</td>
<td>Six LAs attended lecture (assisted students during clicker questions) and held help-room hours to assist students with new conceptual homework</td>
</tr>
<tr>
<td>Sp07</td>
<td>No LAs</td>
</tr>
<tr>
<td>Fa07</td>
<td>No LAs</td>
</tr>
<tr>
<td>Sp08</td>
<td>Seven LAs attended lecture and worked with TAs to co-lead recitation sections (3-4 per week). Principle role of LA is to facilitate student conversations during lecture and to facilitate small group discussions in recitation. Conceptual homework developed in Fall 2006 was integrated into new recitation materials structured for group work.</td>
</tr>
<tr>
<td>Fa08</td>
<td>Plans to use 14 LAs as in Sp08 (47-50 total recitation sections).</td>
</tr>
</tbody>
</table>

*CHEM 1131* (General Chemistry II)
Each week, students attend three 50-minute lectures, one 50-minute recitation session, and one 3-hour laboratory session. (Typically, the recitation and lab occur in the same lab room, in a 4-hour block. Lab follows recitation.) Enrollment typically 250 – 300 students in Fall; 500 - 600 students in Spring. One faculty member teaches one lecture section in Fall; one faculty member teaches two lecture sections in Spring. Graduate TAs lead lab sections, and co-lead recitation sections with LAs.

<table>
<thead>
<tr>
<th>Semester</th>
<th>LA role(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fa06</td>
<td>No LAs</td>
</tr>
<tr>
<td>Sp07</td>
<td>Ten LAs attended lecture, led a total of 18 “learning groups”, and hosted a total of 12 help room hours each week. Of 500 students in course, 150 opted to join a learning group. Principle role of LA was to lead student discussions about new</td>
</tr>
</tbody>
</table>
conceptual homework.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fa07</td>
<td>Five LAs attended lecture and worked with TAs to co-lead recitation sections (3-4 per week). Principle role of LA is to facilitate student conversations during lecture and to facilitate small group discussions in recitation. Conceptual homework developed in Spring 2007 was integrated into new recitation materials structured for group work.</td>
</tr>
<tr>
<td>Sp08</td>
<td>Ten LAs (and one volunteer LA) attended lecture and worked with TAs to co-lead recitation sections (3-4 per week). Principle role of LA is to facilitate student conversations during lecture and to facilitate small group discussions in recitation. Similar to Fa07, but scaled-up to 32 total recitation sections.</td>
</tr>
<tr>
<td>Fa08</td>
<td>No plans to use LAs; all LA resources directed to CHEM 1111.</td>
</tr>
</tbody>
</table>

**Faculty** for AY 2007-2008 in courses involving LAs:

**Fall 2007:**
- Chem 1131 (General Chemistry II): Robert Parson, 5 LAs (1 Returning)
- Chem 3311 (Organic Chemistry I): Rebecca Hoenigman, 2 LAs (both returning—one was paid, one volunteered).

  *This is not listed in the “transformed courses,” because no systematic changes have been made to this course. This instructor was willing to try using LAs in her course. They attended lecture, and the instructor met with them each week and developed materials for use with students. Both LAs were previously CHEM 1131 LAs who had led “learning groups.” Together, we devised a model in which CHEM 3311 students chose to join a group. The first half of the group time was devoted to working on materials provided; during the other half, students could ask their own questions. There is a lot of interest among students and LAs to use LAs in Organic Chemistry. We did collect CLASS data, but it’s in process of being analyzed. No “concept survey” exists for this class.*

**Spring 2008:**
- Chem 1111 (General Chemistry I): Robert Parson, 7 LAs (5 New, 2 Returning)
- Chem 1131 (General Chemistry II): Susan Hendrickson, 10 LAs (8 New, 2 Returning. An additional experienced LA volunteered to help out with one recitation section.)

**Applications** for Fall 2007: 25 new and 6 returning (of the 31 total, 17 had another departmental interest as well).

**Applications** for Spring 2008: 49 new and 5 returning (of the 54 total, 22 had another departmental interest as well).

**Longer term brief summary of LAs in Chemistry:**
We have hired a total of 31 undergraduate LAs since Fall 2006.
Of these, 17 were female = 55%
10 of these total LAs have served two or more semesters.
The rough total number of applicants for chemistry LA positions over 4 semesters is 117.
Findings

LAs

- **Content expertise**

CHEM 1111, General Chemistry I: (Spring 2008) LAs were given the General Chemistry I Concept Survey at the beginning of the semester and again at the end. *(Note: This survey is still in process of revision.)* Due to scheduling issues only 5 of the LAs were able to complete the pre and post concept surveys. The LA data from spring 2008 and TA data from fall 2007 are below for comparison. The TA data is from fall 2007 when all new CHEM 1111 TAs took the concept survey.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pre ave</th>
<th>Post ave</th>
<th>Gain of Ave</th>
<th>Ave Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA (Fall 2007)</td>
<td>17</td>
<td>0.74</td>
<td>0.88</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>LA (Spring 2008)</td>
<td>5</td>
<td>0.71</td>
<td>0.87</td>
<td>0.54</td>
<td>0.70</td>
</tr>
</tbody>
</table>

CHEM 1131, General Chemistry II: (Fall 2007 and Spring 2008) LAs were given the General Chemistry II Concept Survey at the beginning of the semester and again at the end. *(Note: This survey is still in process of revision. Its foci are acid/base and solubility equilibria.)* In the Fall, all 5 LAs took the survey pre and post. Due to scheduling issues in the Spring, only 5 of the LAs were able to complete both the pre and post concept surveys.

<table>
<thead>
<tr>
<th>Semester</th>
<th>N</th>
<th>Pre ave</th>
<th>Post ave</th>
<th>Gain of ave</th>
<th>Ave of Gains (matched)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fa07</td>
<td>5</td>
<td>0.71</td>
<td>0.85</td>
<td>0.48</td>
<td>0.45</td>
</tr>
<tr>
<td>Sp08</td>
<td>7 (pre)</td>
<td>6.9</td>
<td>0.79</td>
<td>0.30</td>
<td>(0.19 for 5 matched LAs)</td>
</tr>
<tr>
<td></td>
<td>10 (post)</td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
</tr>
</tbody>
</table>
Beliefs

CHEM 1111, General Chemistry I: LAs did not take the CLASS survey.

CHEM 1131, General Chemistry II: 10 LAs completed the “pre” survey, and 5 completed the “post” survey. Thus reported shifts are for 5 matched LAs. Overall, these LAs shifted towards a more “expert” view of chemistry. In particular, LAs’ problem solving confidence, sense-making, and atomic-molecular perspectives of chemistry showed large positive shifts. These aspects are emphasized within the recitation materials.

<table>
<thead>
<tr>
<th>Categories</th>
<th>ALL</th>
<th>Number:</th>
<th></th>
<th>LARGE</th>
<th>Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (All 45 Q's with expert response)</td>
<td>favorable</td>
<td>74.2</td>
<td>83.6</td>
<td>9.3</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>8.0</td>
<td>6.2</td>
<td>-1.8</td>
<td>2.0</td>
</tr>
<tr>
<td>All categories (36 Q's that appear in below categories)</td>
<td>favorable</td>
<td>77.8</td>
<td>88.9</td>
<td>11.1</td>
<td>11.1</td>
</tr>
<tr>
<td>Personal Interest</td>
<td>favorable</td>
<td>86.7</td>
<td>90.0</td>
<td>3.3</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>0.0</td>
<td>3.3</td>
<td>3.3</td>
<td>3.0</td>
</tr>
<tr>
<td>Real World Connection</td>
<td>favorable</td>
<td>85.0</td>
<td>90.0</td>
<td>5.0</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>0.0</td>
<td>5.0</td>
<td>5.0</td>
<td>4.5</td>
</tr>
<tr>
<td>Problem Solving General</td>
<td>favorable</td>
<td>86.0</td>
<td>98.0</td>
<td>12.0</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>2.0</td>
<td>0.0</td>
<td>-2.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Problem Solving Confidence</td>
<td>favorable</td>
<td>80.0</td>
<td>100.0</td>
<td>20.0</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Problem Solving Sophistication</td>
<td>favorable</td>
<td>77.1</td>
<td>82.9</td>
<td>5.7</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>8.6</td>
<td>5.7</td>
<td>-2.9</td>
<td>6.3</td>
</tr>
<tr>
<td>SensesMaking/Effort</td>
<td>favorable</td>
<td>77.8</td>
<td>91.1</td>
<td>13.3</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>8.9</td>
<td>2.2</td>
<td>-6.7</td>
<td>2.4</td>
</tr>
<tr>
<td>Conceptual connections</td>
<td>favorable</td>
<td>85.7</td>
<td>91.4</td>
<td>5.7</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>2.9</td>
<td>5.7</td>
<td>2.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Conceptual learning</td>
<td>favorable</td>
<td>62.9</td>
<td>74.3</td>
<td>11.4</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>11.4</td>
<td>11.4</td>
<td>0.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Atomic-Molecular Perspective of Chemistry</td>
<td>favorable</td>
<td>73.3</td>
<td>90.0</td>
<td>16.7</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>unfavorable</td>
<td>10.0</td>
<td>3.3</td>
<td>-6.7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Other LA activities

1) A Noyce fellow studied how TAs and LAs form effective working relationships, which was important in this first year of integrating TAs and LAs in recitation. Her findings are informing our plans for TA and LA training this coming year, and her results are being shared at both the national Noyce conference and the Biennial Conference on Chemical Education.

2) A few “volunteer” LAs have been mentioned in this report. We are finding that several returning LAs who aren’t ready or don’t want to commit to K-12 teaching still want to be involved in the program. They have been willing to commit up to 5-6 hours per week to attend lecture, attend TA/LA meetings, and to fill newly-added recitation slots.
LA-supported courses

Content

CHEM 1111, General Chemistry I: We have been giving the chem 1111 concept survey as a pre/post assessment of learning since spring 2006. The data are below.

<table>
<thead>
<tr>
<th></th>
<th>F06</th>
<th>Sp07</th>
<th>F07</th>
<th>Sp08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre ave</td>
<td>0.41</td>
<td>0.43</td>
<td>0.37</td>
<td>0.36</td>
</tr>
<tr>
<td>Post ave</td>
<td>0.55</td>
<td>0.52</td>
<td>0.50</td>
<td>0.47</td>
</tr>
<tr>
<td>Gain of ave</td>
<td>0.24</td>
<td>0.16</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>N</td>
<td>595</td>
<td>289</td>
<td>674</td>
<td>311</td>
</tr>
</tbody>
</table>

LAs were introduced into CHEM 1111 for the first time in spring of 08. The data above do not demonstrate any increased learning as a result of the use of LAs.

CHEM 1131, General Chemistry II: The General Chemistry II Concept Survey has been in development for three semesters.

- In Spring 2007, only 10 items appeared on both the “pre” and “post” versions. All of those items were related to the topic of acids and bases. That semester, students could choose to join an LA-led learning group. Of the 500 students in the course, about 150 chose to join a group. Thus, only about 30% of students regularly interacted with LAs. Data from the 10 acid/base items are reported below.

- In Fall 2007, the concept survey was revised to only focus on acid/base and solubility equilibria. It is now a 20-item survey. The same version was used “pre” and “post” in both Fall 2007 and Spring 2008. Fall 2007 was the first semester of integrating LAs into recitation, so ALL students were impacted by LAs. We scaled this model up in Spring 2008, so again ALL students were impacted by LAs. Data from these semesters are given below.

<table>
<thead>
<tr>
<th></th>
<th>S07 (10 items; limited use of LAs)</th>
<th>F07 (20 items; LAs in all recitations)</th>
<th>S08 (20 items; LAs in all recitations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre ave</td>
<td>0.50</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>Post ave</td>
<td>0.61</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>Gain of ave</td>
<td>0.22</td>
<td>0.29</td>
<td>0.27</td>
</tr>
<tr>
<td>N</td>
<td>437</td>
<td>183</td>
<td>276</td>
</tr>
</tbody>
</table>
**Attitudes and Beliefs: CLASS-Chemistry Data:**

<table>
<thead>
<tr>
<th>Course and Semester</th>
<th>All Categories</th>
<th>Personal Interest</th>
<th>Real World Connection</th>
<th>Problem-Solving (General)</th>
<th>Sense-Making / Effort</th>
<th>Conceptual Learning</th>
<th>Atomic-Molecular Perspective of Chemistry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Favorable Pre</td>
<td>Shift</td>
<td>Favorable Pre</td>
<td>Shift</td>
<td>Favorable Pre</td>
<td>Shift</td>
<td>Favorable Pre</td>
</tr>
<tr>
<td><strong>CHEM 1111</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fa06 (LAs: limited use) Instructors: 1, 2</td>
<td>59.2</td>
<td>-3.5</td>
<td>57.7</td>
<td>-10.5</td>
<td>62.0</td>
<td>-8.3</td>
<td>62.4</td>
</tr>
<tr>
<td>Fa07 (LAs: none) Instructors: 2, 3</td>
<td>58.0</td>
<td>-2.0</td>
<td>55.7</td>
<td>-6.1</td>
<td>58.9</td>
<td>-3.8</td>
<td>63.5</td>
</tr>
<tr>
<td>Sp08 (LAs: integrated into recitation) Instructor: 4</td>
<td>55.8</td>
<td>-4.6</td>
<td>53.8</td>
<td>-7.3</td>
<td>58.7</td>
<td>-10.8</td>
<td>60.0</td>
</tr>
<tr>
<td><strong>CHEM 1131</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp07 (LAs: limited use) Instructor: 4</td>
<td>62.0</td>
<td>-2.4</td>
<td>55.7</td>
<td>-1.4</td>
<td>62.9</td>
<td>-5.7</td>
<td>65.7</td>
</tr>
<tr>
<td>Fa07 (LAs: integrated into recitation) Instructor: 4</td>
<td>58.5</td>
<td>-0.6</td>
<td>56.0</td>
<td>0.8</td>
<td>57.7</td>
<td>1.5</td>
<td>62.5</td>
</tr>
<tr>
<td>Sp08 (LAs: integrated into recitation) Instructor: 3</td>
<td>58.8</td>
<td>-4.6</td>
<td>54.8</td>
<td>-6.3</td>
<td>59.7</td>
<td>-4.9</td>
<td>63.9</td>
</tr>
</tbody>
</table>

Instructor 1: Several years experience teaching CHEM 1111
Instructor 2: Several years experience teaching CHEM 1111
Instructor 3: Experienced instructor; new to CU in 2007/08, so first time teaching CHEM 1111 / 1131
Instructor 4: Several years experience teaching CHEM 1131; Sp08 was first time teaching CHEM 1111

It is difficult to draw conclusions about the use of Learning Assistants and changes in student attitudes / beliefs given these data. The largest overall negative shifts occurred in semesters when an experienced instructor taught either 1111 or 1131 for the first time (same semesters as full implementation of LAs in recitation). Thus, even though students’ attitudes and beliefs improved in CHEM 1131 from Sp07 to Fa07 (same instructor, increased use of LAs in course), Sp08 CHEM 1111 students had large negative shifts with the same instructor and full use of LAs.
Other Impacts

We have observed several recitation sections run by TAs and LAs. In general, we find that LAs ask better guiding questions and don’t let students “off the hook.” That is, LAs are more apt than TAs to require students to clarify their language and to justify their reasoning. Additional systematic observations are needed to characterize LA-student interactions in Chemistry recitations. Perhaps more immediate assessment tools are needed to measure short-term impact of these LA-student interactions.

Evaluation of department buy-in/challenges

Buy-in

- The department became a member of the Science Education Initiative in Fall 2006. The SEI central funds paid for Chemistry LAs that first year (Fall 2006, Spring 2007). Three faculty members participated in hiring LAs that year.
- In Fall 2008, Chemistry will be using some of its departmental SEI funds to pay for two Learning Assistants in CHEM 1111.
- The Learning Assistant program has been discussed in several General Chemistry Working Group meetings. The SEI departmental liaison led a presentation and discussion of the LA program at a faculty meeting in Spring 2008.
- The SEI post docs (Laurie Langdon and Tom Pentecost) have been heavily involved in hiring and scheduling LAs, as well as co-leading weekly meetings. In Fa07 and Sp08, faculty members co-led LA and LA/TA meetings, and this will continue in Fall 2008. We’re in process of considering structural changes that will make use of LAs sustainable beyond the 5-year Science Education Initiative.
- Two faculty members participated in the August 2007 faculty orientation workshop.
- We are in process of developing course archives for CHEM 1111 and CHEM 1131, including TA/LA/instructor notes for use of recitation materials.
- As of Spring 2008, five different chemistry staff/faculty have been directly involved in teaching courses with LAs.
- In Fall 2007 and Spring 2008, six former Chemistry LAs were hired by the department as undergraduate TAs. Several faculty and staff recognized the outstanding jobs they did in this role (especially compared to a few non-LA undergraduates hired in the same capacity). This has helped to spread positive feedback about the LA program.

Challenges

- LA recruitment and hiring is a large chore which must be done every term. This has largely been done by the two SEI post docs, but as mentioned above, two additional faculty have contributed to the decision-making process in recent semesters. We have also found that a larger pool of applicants does not necessarily translate into an overall better pool of applicants, adding time to the prescreening process.
Baseline data for CHEM 1111 and 1131 (pre Fall 2006) are scarce. In addition, concept survey measurement tools are still in the development / validation stages. So, changes are occurring within the courses at the same time that assessment tools are being developed to measure the effects of those changes. Thus it is difficult to attribute changes in student understanding or attitudes to the use of LAs.

When the gain scores for chemistry courses are compared to other courses, such as physics 1110, the presence of other reforms beyond LAs must be taken into account. The chemistry courses are reformed in the sense of using clickers and LAs in recitation. Work has only recently begun on making the course exams (largest impact on student grades) less algorithmic and more conceptual.

Lack of data that clearly shows positive LA impacts hinders our efforts to convince additional Chemistry faculty to invest the time and effort in using LAs in General Chemistry and other courses. We need to identify other sources of data that will help to make the case.

Lack of departmental funds for LAs. The department is currently committing a portion of its course fees to additional training of its incoming graduate TAs.

References/sources for the various Posttests used: