A LIVELY ELECTRONIC COMPENDIUM OF RESEARCH, NEWS, RESOURCES, AND OPINION

Astronomy Education Review

Volume 7, Aug 2008 - Jan 2009
Issue 2

Effect of Night Laboratories on Learning Objectives for a Nonmajor Astronomy Class

by **Ian C. Jacobi**
Rensselaer Polytechnic Institute
Heidi Jo Newberg
Rensselaer Polytechnic Institute
Darren Broder
Siena College
Rose A. Finn
Siena College
Anthony J. Milano
Rensselaer Polytechnic Institute
Lee A. Newberg
Rensselaer Polytechnic Institute, and Wadsworth Center, New York State
Allan T. Weatherwax
Siena College
Douglas C. B. Whittet
Rensselaer Polytechnic Institute
Received: 08/05/08, Revised: 09/25/08, Posted: 12/16/08

The Astronomy Education Review, Issue 2, Volume 7:66-73, 2009

Abstract

We tested the effectiveness of hands-on nighttime laboratories that challenged student misconceptions, using a new assessment exam to measure learning in a nonmajor introductory astronomy class at Rensselaer Polytechnic Institute. We were able to increase learning at the 8.0 sigma level on one of the Moon phase objectives that was addressed in a cloudy night activity. There is weak evidence of some improvement on a broader range of learning objectives. We show evidence that the overall achievement levels of the four sections of the class are correlated with how much clear weather the sections had for observing even though the learning objectives were addressed primarily in activities that did not require clear skies. We describe our first attempt to cycle the students through different activity stations in an effort to handle 18 students at a time in the laboratories, and the lessons we learned from this.

1. INTRODUCTION

We describe an attempt to use hands-on nighttime astronomy laboratories to improve student learning in an observational astronomy class for nonmajors at Rensselaer Polytechnic Institute. Our challenge was to design meaningful laboratories that could be used at a site that suffers from poor weather, dome seeing, and significant light pollution. We know of no similar study of the effects of a nighttime activity-based learning pedagogy; most studies on the effects of activity-based learning are integrated as a component of the lecture (for example, McCrady & Rice 2008; Adams & Slater 1998; Straits & Wilke 2003). No studies have focused on the effects of activities that include use of a professional observatory setup on learning quality. In fact, Slater (2008) states that although people want to look through telescopes, "It is generally accepted that most undergraduate non–science majors taking an astronomy course spend little to no time actually looking through a telescope." The current focus is more in the direction of Internet-accessible robotic telescopes (see Gould, Dussault, & Sadler 2007, and references therein).

In this article, we assess whether students' learning is improved by moving the activities from the lecture or computer terminal to the real sky. Much of the development and testing of the nighttime laboratories, as well as the analysis of the assessment tests, were done by undergraduate research students.

2. PREPARATIONS

Our primary objective was to determine whether the establishment of a nighttime laboratory component in an introductory astronomy course would improve student achievement. To this end, we created an assessment test in the summer of 2006, which was specifically aligned with the general learning objectives of the introductory astronomy course at Siena College, and the introductory astronomy course, Earth and Sky, at Rensselaer. The assessment test was administered to all students twice—as a pretest and as a posttest—in the Earth and Sky course at Rensselaer Polytechnic Institute during the fall 2006 semester, prior to the creation and implementation of a formal nighttime laboratory component. In this article, we compare the data collected in fall 2006 with data collected after the inclusion of the laboratory component in fall 2007.

We utilized a newly refurbished and automated 16" Boller & Chivens telescope that is mounted atop our science center building. To bring meaningful laboratory experiences to 70–80 students in our nonmajors class, we divided the class into four sections of about 18 students each. Each section met one night every other week. Each section was then subdivided into three groups of six students, and these groups rotated through three activities on each clear night. The stations for the three simultaneous activities were (1) the 16" telescope, (2) the small telescopes on the adjacent rooftop area, and (3) a self-guided naked-eye activity. If the assigned night was cloudy, the students performed a separate set of labs that could be done either in the observatory or in a classroom.

We designed four clear night activities, each of which has three parts (a, b, c), and five activities that could be done in cloudy weather. These activities can be found at http://www.rpi.edu/dept/phys/observatory/labs.html. We discovered that it is difficult to design laboratories that challenge misconceptions or reinforce astronomical concepts using a telescope, particularly an automated telescope with a CCD camera. Although students really like using the big telescope, they are not challenged to confront their misconceptions by using it. Using the smaller telescopes, they are more likely to discover that the sky appears to move; that the size, alignment, and focus of the optics matter; and that the North Star stays in the same place. All of this is handled for them on the automated telescope. It was much easier to teach standard concepts about the phases of the Moon, the paths of the planets, and the effects of the atmosphere in naked-eye and cloudy night laboratories. The important advantage of the 16" telescope was that it was a big draw; all the students wanted to use it.

3. METHODOLOGY

Before this study, the class met in a standard lecture format with homework and exams. Additionally, the students attended a minimum of five "star parties" on clear nights at our campus observatory. The new night laboratories operated two nights per week, and on each night, a team of one graduate student and one undergraduate student ran the activities. Typically, the graduate student was responsible for running the activity on the 16" telescope, the undergraduate student was responsible for running the small-telescope activity, and the naked-eye activities were self-directed.

Even with the significant extra staffing as compared with the previous year, the new laboratories were difficult to implement. It is not easy to convince undergraduate students that they should confront misconceptions and think through the lessons in a group activity. Although most of them enjoy looking at objects through the telescope, Rensselaer students see each assignment as an obstacle to be overcome as quickly as they can, mechanically and without thinking if they can get away with it. The teaching assistants (TAs) were not always prepared to keep students focused on learning the material. The plan was to have two small telescopes on the roof so that three students could be assigned to each, but the students quickly put pressure on the undergraduate TA to set up one telescope and walk them all through the activity together. Students assigned to the naked-eye activity, who were using the same rooftop, often tagged along with the small-telescope group so that there were 8 or 10 students hanging around one small telescope and learning the answers from one another. When the laboratories were run this way, the small-telescope and naked-eye groups always finished ahead of the group in the 16" telescope dome and then had to wait. There was significant pressure on the TAs to get the students out of the lab in well under three hours, especially because the students were spending some of that time just waiting.

The TAs also did not fully grasp the idea of groups cycling through stations. They used the prepared activities as a set of resources from which they improvised a lesson plan. For instance, they discovered that many of the questions in the naked-eye laboratories could be answered without actually looking at the sky; one activity asked the students to use a planisphere and compare it with the sky, but most of the questions could be answered without viewing the sky. So, the students sometimes plowed through a set of naked-eye laboratories on a cloudy night. Because each of the four laboratory sections did a different lab (depending on which groups had good weather) and because the TAs sometimes mixed parts of one lab with those of another lab, it was difficult to keep track of who had done what and when.

4. ANALYSIS AND FINDINGS

Because not all the students remained in the class for the whole the semester, we first removed from the sample those tests given to students who did not take both a pretest and a posttest. Our sample included 82 students in the fall 2006 and 68 students in fall 2007 classes. We compared the performance of the two sets of students on the pretest and concluded that there was no obvious statistical difference between the response patterns.

Figure 1 shows the gain in the fraction of correct student responses in 2006 and 2007 broken down by question. To compute these, we first calculated the fraction of correct responses for each question in the total sample of 150 pretests. The gain in correct answer fraction is the fraction of correct answers on a given year's posttest minus the fraction of correct answers on the combined pretest sample. The most notable results in Figure 1 are the remarkable improvement in Question 11 after the laboratories were introduced, and the decrease in both semesters of the number of students who answered Question 4 correctly. Question 4, which tested students' understanding that resolving power of a telescope depended on both aperture size and wavelength, turned out to be an inaccurate test of student learning; many students learned that the resolving power depends on aperture size, but because they did not also learn that it depends on wavelength (which was confusingly referred to as "color" on the exam), they were drawn toward the wrong answer even though they had in fact learned something. The learning on Question 11 differed between the two semesters at the 8 sigma level. No other question had a difference of more than 2.5 sigma. In addition to the one question that was greatly affected by the learning labs, we found that the fractional change was greater in 2007 in 19 of the remaining 24 questions, suggesting that there may be a smaller effect on learning on a broader set of questions.

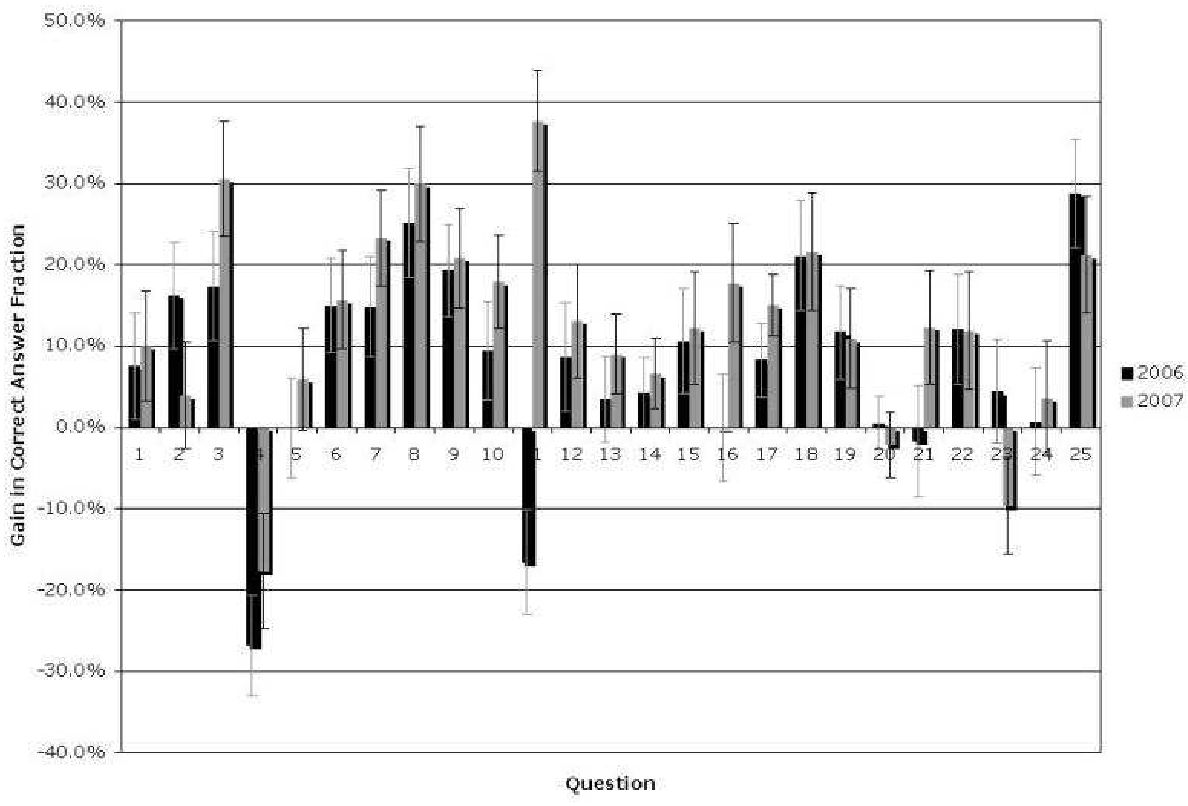

Figure 1. Change in the probability of obtaining the correct answer between the 2006–2007 pretest sample and the posttest sample of a given year, by question number.

Given that the laboratory assignments did not explicitly focus on every question, we examined the change in probability by grouping the questions by their status as having been directly addressed (Questions 1, 3, 6, 9, 11, 12, 16, and 20), indirectly addressed (Questions 4, 7, 10, 13, 15, 22, and 24), or not addressed (Questions 2, 5, 8, 14, 17, 18, 19, 21, 23, and 25) by any lab. Question 21 was classified as *not addressed* because it was addressed only in Lab 4c, which was not completed by any of the sections. Because of the extreme outlier effect that Question 11 was likely to introduce into the findings, we also grouped all *directly addressed* questions together, excepting Question 11.

From the average change in the probability of obtaining a correct answer as categorized by question type, we were able to conclude that there was a statistically significant improvement in all questions addressed by the lab activities, with a 4.1-sigma difference between the improvement in 2006 and the improvement in 2007. If we exclude Question 11, the significance of the improvement in directly addressed learning objectives drops to 1.8 sigma, which is pretty marginal. (For this calculation, we treated Question 11 as if it had never existed. The result would have been slightly stronger if we had tried to evaluate a null statistical model by which the strongest question among the eight directly addressed questions is always excluded.) The significance of the improvement if one considers questions that were addressed directly and indirectly together (excluding Question 11) is 2.5 sigma, which is beginning to be interesting. There was no improvement on questions not addressed by the nighttime laboratories. Overall, this test shows evidence that the nighttime laboratories had an effect, but the statistics were too small to be definitive.

Last, we explored potential differences between the four laboratory sections in 2007, which are labeled Tuesday 1 (T1), Tuesday 2 (T2), Wednesday 1 (W1), and Wednesday 2 (W2). The Tuesday sections were both taught by a team of one graduate and one undergraduate TA, and the Wednesday sections were taught by a different team of one graduate student and one undergraduate TA. Because the weather was different for these sections, they performed somewhat different laboratories. In fact, the second Tuesday section had clear weather every laboratory night, and the second Wednesday section had no clear nights of observing at all. However, the students all did the parts of the labs that addressed the questions on the assessment test, even if the lab was intended for clear nights. We believe that all the sections were exposed to all the material required to do well on the assessment posttest.

We performed a statistical test to determine whether each section was different from the other three by calculating the confidence level at which the 2007 pretest results and the 2007 posttest results for each category in each section were different from the sum of the other three sections. For example, in the 2007 posttest, 10 of 17 students in section W2 got the correct answer for Question 11, and 43 of 48 students in the other three sections (combined) got the correct answer for Question 11. Assuming that W2 was the same as all other sections, we would have expected 13.9 students to get the correct answer and 3.1 students to get the wrong answer. In the other three sections, we expected 39.1 students to get the correct answer and 8.9 students to get the wrong answer. The χ^2 measurement of the difference between the observed and expected values was 7.89. Such a χ^2 value arises by chance with probability only 0.5% (confidence level of 99.5%). The resulting confidence levels that each exam and category in each section are different from the same exam and category for the other three sections are shown in Table 1. The only confidence levels over 90% are in the posttests for the T2 and W2 sections. All cases of a section doing significantly better were in T2, and all cases of a section doing significantly worse were in W2.

Note that the differences were not between sections on different days, which would indicate that the TAs differed in their effectiveness in delivering the material. The results are also not consistently better or worse on the second section taught by each teaching team. The only single factor that we could find that explained all the results was a difference in retention of the material based on the amount of clear weather for each section. The idea that weather was a factor in learning was strengthened by the course evaluations. Although there was a largely neutral response to the laboratories, instructor, and teaching assistants in 2006 and 2007, three of the eight students who wrote comments on the surveys at the end of the 2007 class noted the lack of clear weather. This is consistent with reports from the TAs that students in the second Wednesday class were frustrated with being unable to use the telescopes.

Table 1. Confidence Level of the Detection of a Difference between Each Section and the Other Three

Q#	Pre/Post Test	Tuesday Section 1	Tuesday Section 2	Wednesday Section 1	Wednesday Section 2
11	pretest	56%	48%	84%	86%
	posttest	88%	85%	14%	**99.5%**
Direct	pretest	14%	32%	61%	84%
	posttest	52%	**99.2%**	43%	**98.4%**
Direct w/o 11	pretest	8%	51%	31%	66%
	posttest	78%	**98.0%**	48%	**90.7%**
Indirect	pretest	32%	62%	32%	60%
	posttest	58%	43%	15%	5%
Both w/o 11	pretest	28%	73%	43%	79%
	posttest	85%	**95.8%**	45%	74%
Unaddressed	pretest	30%	70%	21%	30%
	posttest	57%	6%	75%	**92.9%**
All Questions	pretest	10%	3%	62%	68%
	posttest	25%	**91.3%**	75%	**98.6%**

Note: Red numbers indicate that the section scored significantly worse than the other three, and green numbers indicate that the section scored significantly better. The only comparisons that differ at more than 90% confidence are for Tuesday Section 2, which in all cases did better than expected, and for Wednesday Section 2, which in all cases did more poorly than expected.

It should be stressed, though, that our experiment was not designed to test this hypothesis, and the finding is therefore the result of multiple hypothesis testing. One fact that causes us to question the validity of this hypothesis is that the pretest results for Question 11 in section W2 were low compared with the other sections. Further studies should be designed to test the relationship between clear weather and learning.

5. DISCUSSION AND CONCLUSIONS

Using a 25-question assessment test developed for this study, we tested the effectiveness of nighttime laboratories (http://www.rpi.edu/dept/phys/observatory/labs.html) on learning in a nonmajor introductory astronomy class at Rensselaer Polytechnic Institute. Of the 25 questions, 8 were directly addressed by the laboratories, 7 were indirectly addressed by the laboratories, and 10 were not addressed in the laboratory but were addressed in the lecture portion of the class, which remained unchanged between 2006 and 2007. We suggest that the fourth question on our assessment test should be discarded or reworded because it generated ambiguous results that were difficult to interpret.

For Question 11, a geometrical question that tested whether students know that the full Moon cannot be seen at noon, there was an enormous (8 sigma) improvement in learning after the introduction of the nighttime laboratories. On the pretest, 45% of students in the combined 2006 and 2007 classes chose the correct answer. In 2006, only 28% of the class had the right answer on the posttest, but in 2007, 82% of the class answered this question correctly. We attribute this improvement to the cloudy activity, "Light and Shadow in the Solar System," which was introduced to the 2007 class.

There was general improvement in the learning of the other 14 questions addressed by the nighttime laboratories, but at low significance. There was no overall change in learning on questions that were not addressed by the night laboratories; simply introducing new laboratories did not stimulate the students to study the general course material at a higher level.

The one section out of four that showed slightly less learning in all categories in 2007 was the section that never had a clear night. The section with the best weather showed the highest learning achievement. These results are suggestive but are subject to a critique of multiple hypothesis testing; furthermore, the results change somewhat with different analysis techniques, so they should be specifically tested for in future learning studies. If we assume that our result is correct, we conclude that the atmosphere in which the material is presented has a strong effect on whether the students learn the material.

We pioneered the idea of laboratory rotations in the nighttime laboratories and encountered significant obstacles in the required nighttime staffing level, the amount of time required from the students taking the course, timing the switch between stations, and coordinating the chaos of running a different laboratory every night. We discovered that undergraduate and first-year graduate student TAs found running the laboratories extremely challenging. The students were not prepared to be out until 10:00 or 11:00 p.m., working in the observatory on their assigned nights. It was difficult to time all three activities so that they took the same amount of time; any equipment difficulties on the 16" telescope put that portion of the laboratory way behind schedule. Because each section had a different number of clear nights, and we needed both a clear and a cloudy laboratory ready for any given night, it was challenging to prepare to teach these classes. Hemenway et al. (2002) found that results may vary between the first and second semesters during which a class has implemented activities-based learning, with the second semester of activity-based learning showing more significant improvement in both astronomy content and the course survey. Given the administrative difficulties we encountered when introducing this new curriculum, we would not be surprised to find that learning and course survey results improve the second time around for this set of activities. It is clear that more effort should be placed on training TAs who run the night laboratories.

In the future, we plan to reduce the stress of the nighttime laboratories on both the TAs and the students while keeping those activities that have been shown to significantly improve student learning. Instead of using rotations through three stations, we plan to schedule smaller groups to be in the observatory for shorter periods. During the time they are in the observatory, students will work with a TA in a smaller group so that they will spend less time waiting. It turned out to be unrealistic to expect nonmajors to quickly learn how to operate the small telescopes, so laboratories that require this skill will be modified or eliminated. With shorter laboratories (which the students will attend every week rather than every other week), there will be less variation in the amount of clear weather that each section experiences.

A longer version of this article, including a fuller description of the project, a copy of the assessment test, the numerical data, and a more complete description of the statistical techniques, is available on the LANL preprint server (http://arxiv.org/abs/0809.3817).

This work was supported by NSF Grant No. DUE 05-11340 and the NASA/NY Space Grant.

References

Adams, J. P., & Slater, T. F. 1998, "Using Action Research to Bring the Large Lecture Course Down to Size," *Journal of College Science Teaching*, 28(2), 87.

Gould, R., Dussault, M., & Sadler, P. 2007, "What's Educational about Online Telescopes? Evaluating 10 Years of MicroObservatory," *Astronomy Education Review*, 5(2), 127. http://aer.noao.edu/cgi-bin/article.pl?id=217.

Hemenway, M. K., Straits, W. J., Wilke, R. R., & Hufnagel B. 2002, "Educational Research in an Introductory Astronomy Course," *Innovative Higher Education*, 26(4), 271.

McCrady, N., & Rice, E. 2008, "Development and Implementation of a Lab Course for Introductory Astronomy," *Astronomy Education Review*, 7(1), 13. http://aer.noao.edu/cgi-bin/article.pl?id=269.

Slater, T. 2008, "The First Big Wave of Astronomy Education Research Dissertations and Some Directions for Future Research Efforts," *Astronomy Education Review*, 7(1), 1. http://aer.noao.edu/cgi-bin/article.pl?id=261.

Straits, W. J., & Wilke, R. R. 2003, "Activities-Based Astronomy: An Evaluation of an Instructor's First Attempt and Its Impact on Student Characteristics," *Astronomy Education Review*, 2(1), 46. http://aer.noao.edu/cgi-bin/article.pl?id=44.