Development of Scientific Abilities in a Large Class

Sahana Murthy and Eugenia Etkina

Graduate School of Education, Rutgers University, New Brunswick, NJ 08904

Abstract. This paper describes our instructional and research efforts to help students in a large-enrollment (450 students) introductory laboratory course develop abilities used by practicing scientists. We focus on the ability to design an experimental investigation. We provide sample tasks, scoring rubrics and evidence of student improvement.

INTRODUCTION

The practice of science and engineering requires not only content knowledge but also other specific abilities that our graduates have to develop to be successful in the future [1,2]. These scientific abilities include formulating questions, designing and conducting experiments, collecting, representing and analyzing data, modeling, testing hypotheses and solving complex, ill-defined problems [3]. They form as the result of training and practice. We believe that a physics class is an excellent place for students to learn these abilities [4]. However, in terms of helping students develop scientific abilities, most attempts are in middle-school [5,6] with fewer examples at college level. Some college examples are the SCALE-UP project [7] and the Workshop Physics [8] project. These projects have been implemented in specially designed classrooms or in smaller classes. There have been even fewer attempts to help students develop scientific abilities in large-enrollment courses with a traditional structure. This paper describes our instructional and research efforts to help students develop experimentation-related abilities in a large-enrollment (450 students) introductory laboratory course.

EXPERIMENTAL DESIGN TASKS

To help our students develop scientific abilities we devised laboratory tasks where students design experiments to test a hypothesis or to solve a problem. These tasks have a number of important features. They are open-ended and some of the information required to solve them has to be obtained by different means, often by performing additional experiments or by making informed estimates. Students have to design and describe their own procedure to solve the task. The tasks encourage divergent thinking, as students need to come up with at least two independent methods to solve a problem. A task where students have to solve a practical problem is shown below:

Sample design task: Design experiments to determine the thickness of a strand of your hair using two independent methods. One of the methods must involve ideas about diffraction. Available equipment includes laser pointer, ruler, paper, and holder for strand of hair, Vernier calipers. For each method write in your lab report:

a) Give an outline of your experimental design.
b) Draw a labeled diagram of your set-up.
c) Write the mathematical procedure you will use.
d) Write how you will measure the physical quantities you need to determine the thickness.
e) Perform the experiment and record your measurements in a table.
f) Calculate the thickness, based on your procedure and measurements.
g) Identify sources of experimental uncertainty. Write the steps you can take to minimize them?
h) Compare the two values you obtained for the thickness of the hair in two experiments. Describe possible reasons for the difference.

SCIENTIFIC ABILITY RUBRICS

Together with the Rutgers University Physics and Astronomy Education (PAER) group members, we developed scoring rubrics to evaluate students' lab reports. The rubrics contain descriptors for individual scientific abilities on a scale of 0 to 3. Table 1 shows some elements of the rubric that we used to score students' lab reports for experiment design tasks. The complete rubrics, which contain all the scientific abilities are available on the group's website [9]. The rubrics were extensively tested for inter-rater reliability. Over a period of two months, we scored many student write-ups and revised our rubrics iteratively until we achieved an agreement of 90-95% in the scores.

We also used the rubrics for another purpose. As seen in the sample, we divided the design problem into sub-tasks (parts a-h in the sample), so that each sub-task reflected a scientific ability that we would like students to develop. We structured the tasks to achieve a correspondence between the guidelines to the students and the abilities in the rubrics. Thus our design tasks also underwent an iterative process of revision along with the rubrics. It is important to note that although we provided detailed guidelines we did not provide a recipe for solving the problem. The guidelines provided a template for all experiments of this type. Consequently, the rubrics helped to develop new design tasks.

There is another important but subtle aspect of the rubrics. Although the same rubrics can be used for different design tasks, they are not free of physics content. The ability to design a reliable experiment to solve the problem encompasses physics content knowledge. In the same vein, most sub-tasks for different design problems look similar, yet the process of devising the correct procedure to solve the problem involves a thorough understanding of the physics concepts involved.

IMPLEMENTION OF DESIGN TASKS

These tasks were implemented in the second semester of a laboratory course that accompanied an introductory physics lecture-recitation course for science majors. Almost all students who took the laboratory course were enrolled in the lecture-recitation course. In the first semester of the lab course, students performed non-cookbook style open-ended experiments, some of them being design tasks. However, in the first semester the tasks did not have guidelines that reflected scientific abilities.

TABLE 1. . Some elements of a scoring rubric used for a problem-solving design experiment

Ability \ Score	0	1	2	3
1. Is able to design a reliable experiment that solves the problem	The experiment does not solve the problem.	The experiment attempts to solve the problem but due to the nature of the design the data will not lead to an accurate solution.	The experiment attempts to solve the problem but due to the nature of the design there is a moderate chance the data will not lead to an accurate solution.	The experiment solves the problem and has a high likelihood of producing data that will lead to a reliable solution.
2. Is able to choose a productive mathematical procedure for solving a particular experimental problem	Mathematical procedure is either missing, or the equations written down are irrelevant to the experimental design	A mathematical procedure is described, but it is incomplete, due to which the final answer cannot be calculated.	Correct and complete mathematical procedure is described but an error is made in the numerical calculations.	Mathematical procedure is fully consistent with the design. All quantities are calculated correctly. Final answer is meaningful.
3. Is able to communicate the details of an experimental procedure clearly and completely	Diagrams are missing and/or experimental procedure is missing or extremely vague.	Diagrams are present but unclear and/or experimental procedure is present but important details are missing.	Diagrams and/or experimental procedure are present but with minor omissions or vague details.	Diagrams and/or experimental procedure are clear and complete.
4. Is able to evaluate specifically how experimental uncertainties may affect the data	No attempt is made to evaluate experimental uncertainties.	An attempt is made to evaluate experimental uncertainties, but most are missing, described vaguely, or incorrect.	Most experimental uncertainties are evaluated correctly, though a few contain minor errors, inconsistencies, or omissions.	All experimental uncertainties are correctly evaluated.

During the semester the project was implemented, there were 20 lab sections, each with about 25 students. There were nine Teaching Assistants (TA's): five were first year TAs with a first language other than English, and four were engineering graduate students. None of the TAs was involved in PER activities. We collaborated closely with the course coordinator who was greatly instrumental in making possible the implementation of our design tasks. One of the authors (SM), in association with the course coordinator, led TA training sessions. During an hour-long weekly meeting, the TAs went through a mock process of performing the activities. They described how they would design the experiment, what they would measure and how they would solve the task.

FINDINGS AND DISCUSSION

We found that in the initial weeks, students had difficulties approaching design tasks. Their lab reports received low scores on scientific abilities, as seen in Figure 1(left). We observed students in the laboratory and noticed negative attitudes and frequent complains that the tasks were difficult. As the semester progressed the students became more enthusiastic about these activities. They spent more time discussing in their groups how to come up with different methods to solve the task. The quality of their write-ups changed. The reports began to resemble experimental reports of practicing scientists. See Figure 1(right) which is a student response to the sample design task. These observations are similar to those of Zou, who implemented design tasks in small classes [10].

We selected four laboratory sections taught by different TAs to sample students' lab write-ups. We included 35 students in the sample. The students were also randomly distributed from 17 recitation sections of the lecture-recitation course. The average final exam score for the sample was 78.3, class average was 75 with a standard deviation of 16. We scored reports of the students chosen above on different scientific abilities. As seen in the bar charts in Figure 2, there was an improvement in the scores for certain scientific abilities. A closer examination using Chi-square analysis revealed that the improvements were statistically significant. Examining each ability we found contributors to the significant changes (see Table 3): For ability 1 there was a significant change in ranking 1; for ability 2 - a significant change ranking 3; for ability 3: - significant changes in rankings 0 and 3; for ability 4 – a significant improvement with no significant contributing cells. We also found no significant χ^2 values when the course grades were compared to a particular ability ranking. Possibly, our data sample was not large enough to see the significant differences.

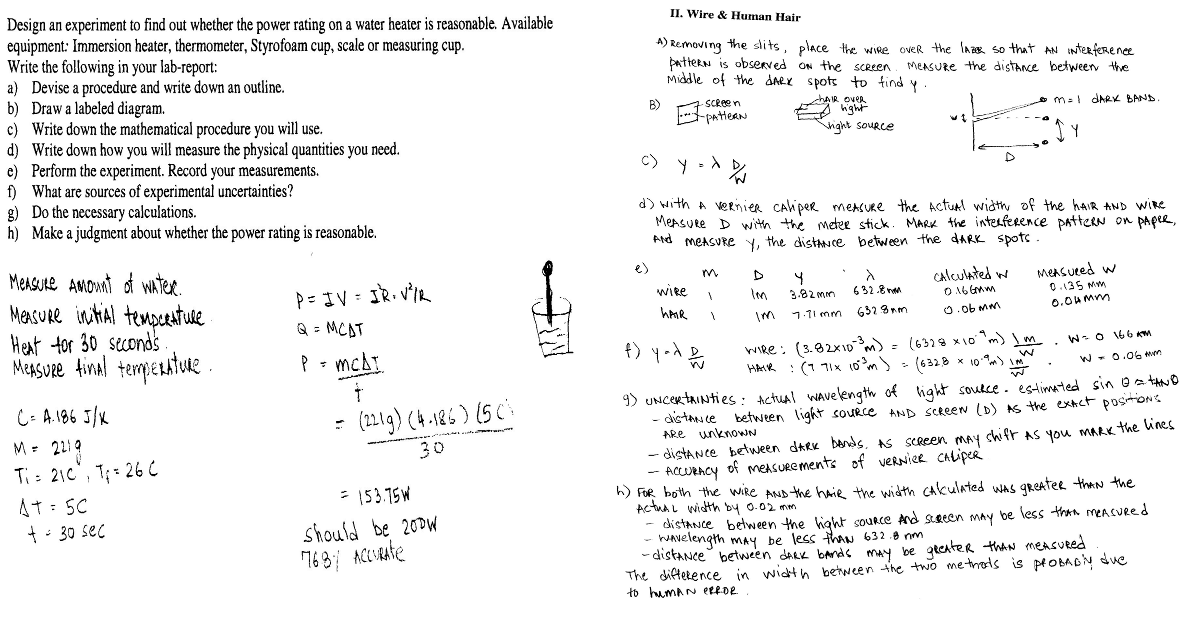
Design an experiment to find out whether the power rating on a water heater is reasonable. Available equipment: Immersion heater, thermometer, Styrofoam cup, scale or measuring cup.
Write the following in your lab-report:
a) Devise a procedure and write down an outline.
b) Draw a labeled diagram.
c) Write down the mathematical procedure you will use.
d) Write down how you will measure the physical quantities you need.
e) Perform the experiment. Record your measurements.
f) What are sources of experimental uncertainties?
g) Do the necessary calculations.
h) Make a judgment about whether the power rating is reasonable.

FIGURE 1. Sample lab reports from students.
Left: From week 3. Scores given are as follows: Ability1 -- 2, Ability 2 -- 2, Ability 3 --1, Ability 4 – 0.
Right: From week 10, design task is in the text as sample design task. Scores given are as follows: Ability1 -- 3, Ability 2 -- 3, Ability 3 --2, Ability 4 – 2.

TABLE 2. Chi-square analysis		
Ability	χ^2 ($\chi^2_{critical\ value}$=7.82)	Standardized residual, R (R> |2| is a significant contributor to the χ^2 value)
1	17.03	$R_{i,1}$=2.23; R_{f1}=-2.23
2	17.73	R_{i3}=-2.23; R_{f3}=2.23
3	25.12	R_{i0}= 2.55; R_{f0}=-2.55 R_{i3}=-2.37; R_{i3}=2.37
4	15.32	No cells were significant contributors.

What can we conclude from our results? The first conclusion is that it is possible to implement and assess open-ended tasks even in large-enrollment classes, without totally revising the course and having highly trained teachers/TAs. The second conclusion is that students' scientific abilities as measured by our rubrics improved significantly. There could be multiple explanations for this effect, one of which is that the sub-tasks provided guidance. We favor this explanation because for the fourth ability there were no guidelines (see sample task) and there were no significant standardized residuals found. We believe there is a need for a controlled experiment to test this hypothesis. Another conclusion is that the work on the assessment rubrics changed the tasks that we gave to the students, the wording of which was revised to match the rubric's criteria. Descriptors in the rubrics served as goals for writing design tasks. Revisions of the rubrics led to the revisions of the tasks.

We plan to make the scoring rubrics available to the students in each lab next semester. This can serve the purpose of providing both guidelines and self-assessment, which is regarded to be a very good form of assessment [11]. We also intend to train TAs on how to use the scoring rubrics, so that they can be used as a part of grading. More tasks from different areas of physics can be found on the group's website [9]. We hope physics teachers in different settings can use them along with the scoring rubrics.

ACKNOWLEDGEMENTS

We thank S. Brahmia, D. Brookes, M. Gentile, H. Lopez, M. Milner, D. Rosengrant, Y. Timofeeva, A. Van Heuvelen, and A. Warren for helping in the development of rubrics, and M. Lawrence for help in preparing the paper. We are grateful to G. Alba and R. Fiorillo for support in implementing design tasks. The project was supported by NSF grant DUE 024178.

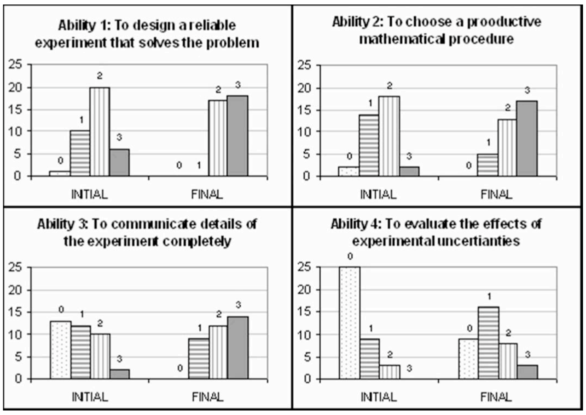

FIGURE 2. Scores of 35 students on the scientific abilities listed in Table 1. INITIAL refers to week 3 and FINAL refers to week 10 in the semester.

REFERENCES

1. See the ABET home page at http://www.abet.org/abethomepage.html.
2. Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology," NSF Directorate for EHR Review of Undergraduate Education, May, 1996.
3. Schunn, C. and Anderson, J., "Acquiring expertise in science: Exploration of what, when and how," in *Designing for science: Implications from everyday, classroom and professional settings* edited by K. Crowley, C. D. Schunn, and T. Okada, Mahwah, NJ: Erlbaum. 2001, pp. 351-392.
4. Etkina, E., Developing scientific abilities in a large-enrollment physics course. Winter AAPT, 2004.
5. Black, P. J. "Can pupils design their own experiments?", *P.E.E. Conference,* Nankai University Tianjin, April 1990.
6. Germann, P. J., Aram, R. and Burke, G., "Identifying Patterns and Relationships among the Responses of Seventh-Grade Students to the Science Process Skill of Designing Experiments", *Journal of Research in Science Teaching,* **33**(1) (1996).
7. See the SCALE-UP project's website at: http://www.ncsu.edu/per/scaleup.html
8. See the Workshop Physics website at: http://physics.dickinson.edu/~wp_web/WP_homepage.html
9. See Rutgers Physics and Astronomy Education Research group's website: http://paer.rutgers.edu/ScientificAbilities/
10. Zou, X., Etkina, E. and Van Heuvelen, A., "Experiment designing and hypothesis testing in the introductory laboratory," National AAPT Conference, Rochester, New York, July 2001.
11. Black, P. and William, D., "Assessment in Classroom Learning", *Assessment in education*, **5**(1), pp. 7-75 (1998).