

Peer-assessment of Homework Using Rubrics

Sahana Murthy

Massachusetts Institute of Technology, 77 Mass. Ave., Cambridge, MA 02139

Abstract. I have implemented a peer-assessment system in an introductory physics course, where students assess each other's homework. Students are provided with descriptive rubrics to guide them through the process. In this paper I describe the implementation of the peer-assessment process, discuss the role of rubrics, present data of agreement of students' assessment with the instructor's, and show evidence of student improvement in evaluation abilities.

Keywords: Rubrics, assessment, peer assessment, evaluation.

PACS: 01.40.Fk, 01.30.lb, 01.40.gb

INTRODUCTION

An assessment system that supports learning involves student participation in productive activity, distribution of student effort evenly across topics and weeks, communication of clear expectations by the instructor, and detailed, frequent and quick feedback that reinforces learning goals [1]. One way to meet the above goals is to have students systematically and consistently assess the work of their peers. Peer-assessment also engages students in the process of evaluating scientific information, which is often an important part of students' future professional careers [2]. In addition, peer-assessment could create a sense of community in the classroom and a shared ownership of the learning process. Instances of peer-assessment have previously been reported in middle school classrooms [3] and college-level courses [1] with favorable results.

In this paper I describe instructional and research efforts to implement a peer-assessment system in an introductory physics class in which students assess each other's homework. Students were provided with assessment rubrics to guide them through this process, and to incorporate descriptive, criterion-based feedback. In this preliminary study, I address two questions: how consistent is students' assessment with the instructor's (author of this paper), and do they develop and improve upon evaluation abilities?

COURSE DETAILS

The peer-assessment process was implemented in a two-semester calculus-based introductory physics course. The course is part of MIT's Experimental

Study Group [4], an alternate academic program that offers highly interactive, small group learning in the core first-year subjects within a community-based setting. There were two sections, each of about 10 students, consisting mainly of engineering majors. Each section met for 4 hours a week. Instruction included in-class collaborative problem solving, some ISLE-style observational and testing experiments [5], visualizations [6], and class discussions. Throughout the course, there was a focus on developing and using scientific abilities such as multiple representation, experimental testing and evaluation [7].

Homework details

Homework was assigned on a weekly basis. Students wrote their solutions and submitted them in class a week later. Homework contained various PER-based questions, such as multiple-representation tasks [8], ranking tasks [9], convince-your-friend tasks [8], and evaluation tasks [10]. Longer analytic questions had parts that focused on scientific abilities. A sample question from kinematics is shown below:

Example. A car and a motorcycle travel on a straight highway. The motorcycle is initially at a distance d behind the car and moves at the same velocity as the car. The motorcycle starts to pass the car by speeding up at a constant acceleration a_m . When it is side by side with the car, it stops accelerating and is travels at twice the velocity of the car. Your goal is to determine the distance traveled by the motorcycle while it is accelerating.

- a) Construct a position-time and a velocity-time graph for the two vehicles.

- b) Determine the distance the motorcycle traveled while it was accelerating.
- c) Perform at least two checks to see if your result is reasonable.

ELEMENTS OF THE PEER-ASSESSMENT PROCESS

Rubrics

The backbone of the peer-assessment process was the assessment rubrics. The rubrics contained detailed but general descriptors of criteria to assess problems. These criteria include: physics content, relevant representations, modeling the situation, problem-solving strategy and reasonableness of answer. The criteria were chosen based on problem-solving literature in physics and science education [such as 12, 13]. For each criterion, the rubrics contained a scoring scheme on a scale of 0-3 for different levels of performance, 3 being the desired level. The format of the rubrics is based on those described in [7]. Students used the same assessment rubrics throughout the year. A portion of the rubrics is shown in Table 1.

In addition, students were also provided with a taxonomy that described what the criteria in the general rubrics were, for a specific problem. For example, the taxonomy for the problem solving strategy in the example above includes choosing a coordinate system, identifying physical quantities with respect to the coordinate system, and equating the position of the two vehicles at the time of their meeting.

Timeline of the assessment process

Following students' submission of their homework, there was a 10-minute class discussion of assessment criteria for selected questions. Then each student was randomly assigned another student's homework for assessment. Students also took home solutions to the homework questions, rubrics, taxonomies and a score-sheet. Fig. 1 shows a sample score-sheet filled out by a student-grader. The score-sheet contained a blank grid in which students wrote the rubric score and comments for relevant criteria for every homework question. If they assigned a less than perfect score, they had to write a comment as to why the particular score was assigned. For example, see the entry under problem 1, 'Relevant representations' in Fig. 1. The student-grader has assigned a score of 2, and has commented that labels on forces are missing.

Students returned the assessed homework in 3-4 days to the instructor who then checked student assessment and made necessary corrections on the score-sheet. See for example, the circled entries (initialed by the instructor) under problem 4 in the score-sheet in Fig. 1. If a score-sheet contained too many errors, the instructor met with the student-grader and discussed the problems. Such meetings occurred more frequently and with more students in the initial weeks, but as students' assessment became more consistent, the need for these discussions reduced. Students got their homework with rubric scores and comments in 1-2 days after they handed in filled out score-sheets. The entire process from the time students handed in their homework to the time they got back their assessed homework took no longer than a week.

TABLE 1. Portion of assessment rubrics showing two criteria.

CRITERIA	0: MISSING	1: INADEQUATE	2: NEEDS IMPROVEMENT	3: ADEQUATE
Modeling the situation Is able to construct a useful model of the situation in the problem	No attempt is made to model the situation.	An attempt is made to model the situation but details are missing or unclear.	An attempt is made to model the situation. Some assumptions are stated, but there is an error or one important aspect missing.	The model fits the given situation. All assumptions about objects and processes are clearly stated.
Reasonableness of answer Is able to evaluate if the answer obtained is reasonable.	The answer is not reasonable and no attempt is made to evaluate its reasonableness.	The answer is reasonable, but student makes no attempt to evaluate why it is so. OR: The answer may not be reasonable and student recognizes that, but does not attempt to analyze what makes it so.	The answer is reasonable, and student makes an attempt to evaluate it, but the evaluation is incorrect or incomplete. OR: The answer may not be reasonable but student recognizes that and tries to analyze it. But the analysis is flawed or incomplete.	The answer is reasonable, and student evaluates why it is so (e.g. by limiting cases or dimensional analysis). OR: The answer may not be reasonable. Student correctly analyzes what makes it so.

8.01 PROBLEM SET 6 SCORING SHEET

STUDENT NAME: [REDACTED] DATE: 11-6-06
 GRADER NAME: [REDACTED] GRADER SIGNATURE: [REDACTED]

	Problem 1.	Problem 2.	Problem 3.	Problem 4.	Problem 5.	Problem 6.
Physics content	3	1 Mixed ideas of energy and momentum	3	② → 3 SH	3	3
Relevant Representations	2 Labels on force diagram	3	N/A	0 Missing	2 No axes	3
Modeling the problem	2 Wrong coord - inate choice	2 Energy not conserved always	3	2 no assumptions	2	3
Problem-solving strategy	3	2	2 Rethink reasoning	③ → 2 SH CHECK SIGNS	3	3
Reasonableness of answer	2 Check?	1 Wrong. No check	3	3	3	2 check dimensions

FIGURE 1. Score-sheet filled out by student-grader

Connection between peer-assessment and grades

The grade a student received on a homework was computed by adding the checked rubric scores from the score-sheet. The homework grades were worth 20% of the course grade. I considered reserving a portion of the course grade for peer-assessment, but I did not encounter many problems in terms of students not taking peer-assessment seriously. In the end, no extra points were assigned for the peer-assessment part of the homework, but a student had to return a classmate's assessed homework in order to receive credit for his/her own homework.

RESULTS

Consistency between students' and instructor's assessment

When students returned the assessed homework and score-sheet, the instructor checked their entries for

rubric scores. Essentially, a student's homework was scored again, using the same rubrics. Discrepancies between student-assigned scores and instructor-assigned scores were noted. The left graph in Fig. 2 shows first semester (mechanics) data for the average number of such discrepancies per week. The total number of rubric scores in a given week was between 30 and 40, depending on the number of homework questions. We see that by week 4, the average number of discrepancies have stabilized to about 2 per week (about 7%). In the right graph in Fig. 2 we see the percentage of students whose assessments had more than 10% discrepancy with the instructor's assessment. Again we see that this percentage steadily decreases initially, and then stabilizes around week 4. A similar documentation (graph not shown) of assessment discrepancies from the second semester (electricity and magnetism) shows that students' assessment was consistent with the instructor's.

Development of evaluation abilities

Evaluation ability is defined as making a judgment about information based on specific criteria [7] and strategies. Since the process of peer-assessment involves judging scientific information based on specific criteria in the rubrics, a reasonable question to ask is if students' evaluation abilities change during the course of the semester. That is, can students learn strategies used by practicing physicists, such as dimensional analysis and special case analysis to evaluate a solution to a problem? To explore this, exam questions were assigned in which students had to evaluate a proposed solution to a given problem. The solution had errors in physics content, incorrect assumptions, numerical errors and so on. Students had to evaluate the proposed solution using different strategies, discuss whether the solution satisfied certain criteria and identify errors.

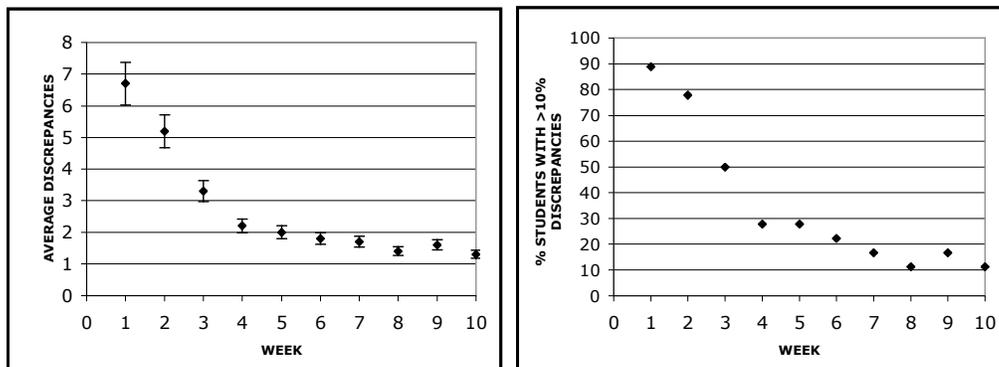


FIGURE 2. LEFT: Average number of assessment discrepancies per student per week. RIGHT: Percentage of students with 10% or more assessment discrepancies per week

Students' responses were scored using scientific abilities rubrics developed by the Rutgers PAER group [11]. (These are not the same as the rubrics students used for peer-assessment). Students' average scores on three evaluation strategies are shown in Table 2.

Strategy	Exam 1	Exam 2	Final
Use of dimensional analysis	2.5 (0.95)	2.8 (0.62)	2.9 (0.51)
Use of special-case analysis	1.4 (1.04)	1.8 (0.98)	2.6 (0.65)
Identification of errors in modeling	1.6 (1.18)	2.0 (1.10)	2.2 (1.09)

Students quickly mastered the strategy of dimensional analysis to evaluate a solution. The strategy of using special case analysis took longer to be learned. Students struggled the most with identifying errors in assumptions and modeling a situation (such as, is it constant velocity or constant acceleration motion), but they improved upon this as the semester progressed.

DISCUSSION AND SUMMARY

Preliminary data from the first semester show that students do learn to systematically assess physics problem solutions using given criteria. In our study, students' assessment could be considered to be consistent with the instructor's after about four weeks. Students' evaluation abilities improved over the course of the semester. Some factors which contributed to the success of the peer-assessment process are: the use of rubrics which helped clearly communicate learning and performance goals and incorporated criterion-based descriptive feedback; a weekly schedule that encouraged quick feedback; dedicated class time to discuss details of assessment; and an effort by the instructor to develop among students a sense of the importance of evaluation abilities.

Implementing the peer-assessment process was not without challenges. The initial investment of time in writing the rubrics was large. Substantial effort was required in training students for the first 3-5 weeks. The overall time spent by the instructor was comparable to what one would spend grading homework traditionally. In terms of attitudes, most students accepted peer-assessment as part of the course. In class, the instructor repeatedly discussed the rationale behind peer-assessment and invited students to share their reactions. On the positive side, some students thought that peer-assessment helped them realize what they did or did not understand. On the other hand, one student, who mainly focused on the

grading aspect, repeatedly expressed her discomfort. She commented on the mid-term evaluation that it was the teacher's job to grade, not the students'.

In terms of further work, one question to study is if there students' assessment work in a given content area is related to their conceptual understanding. Another open question is to understand what frame students are in when they are involved in assessment tasks: do they see it as a grading task, do they reflect on their understanding or are they looking to learn something new? A practical problem that needs to be addressed is how to scale up the peer-assessment process to larger classes. These are the subjects of present and future studies. Currently, results from this study are being used to develop and implement a peer-assessment process in two other classes at the Experimental Study Group.

ACKNOWLEDGMENTS

I would like to thank all members of the Experimental Study Group at MIT.

REFERENCES

1. Gibbs, G. and Simpson, C. "Does your assessment support your students' learning?" *Journal of Learning and Teaching in Higher Education*, **1**, 1 (2003).
2. Boud, D, "The role of self-assessment in student grading", *Assessment and Evaluation in Higher Education*, **14**, 20-30 (1989).
3. P. Sadler and E. Good, "The impact of self- and peer-assessment of student learning", *Educational Assessment* **11**(1), 1-31 (2006).
4. <http://web.mit.edu/esg/www/>
5. E. Etkina, A. Van Heuvelen, D. T. Brookes, and D. Mills, *Physics Teacher* **40**, 351 (2002).
6. <http://web.mit.edu/8.02t/www/802TEAL3D/>
7. E. Etkina, A. Van Heuvelen, S. White-Brahmia, D. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and A. Warren. "Scientific abilities and their assessment." *Phys. Rev. ST Phys. Educ. Res.* **2** (2), 020103 (2006).
8. A. Van Heuvelen and E. Etkina, *The Physics Active Learning Guide*, San Francisco: Pearson Education, 2006.
9. T. L. O'Kuma, D. P. Maloney and C. J. Hieggelke, *Ranking task exercises in physics*, Benjamin Cummings, 2003.
10. For a description of these tasks, see <http://paer.rutgers.edu/ScientificAbilities/Formative+Assessment+Tasks/default.aspx>
11. <http://paer.rutgers.edu/ScientificAbilities/Rubrics/default.aspx>, specifically Rubric I.
12. R.E. Mayer, *Thinking, Problem solving, Cognition* (2nd ed.), New York: Freeman & Co., 1992.
13. A. Van Heuvelen. "Learning to think like a physicist: A review of research-based strategies," *Am. J. Phys.* **59**, 891-897 (1991).