

# Effect of peer-review on development of students' problem-solving abilities

Taylor Crouch<sup>1</sup> and J. Christopher Moore<sup>1</sup>

<sup>1</sup>*Department of Physics, University of Nebraska Omaha, 6001 Dodge St., Omaha, NE, 68182, USA*

We have investigated the effect of peer-review on the development of students' problem-solving abilities in an introductory physics course. Specifically, we report the results of a multiple-group pre/post-test quasi-experiment comparing two groups receiving different treatments with respect to activities completed after cooperative group problem-solving (CGPS) sessions. Both the treatment and control groups received identical instruction in lecture and CGPS discussion sections. Individual student problem solutions submitted on-line were evaluated via a validated problem-solving process rubric that was supplied to all participants at the beginning of the course. The treatment group completed rubric-based peer-review on-line for three randomly selected students, whereas the control group did not. Reliability-corrected analysis of covariance showed that the treatment group demonstrated greater improvement in problem-solving process compared to the control over the four-week study period.

## I. INTRODUCTION

Problem-solving is recognized as an important component of the introductory physics course and science courses in general [1]. In particular, physics education research teams have identified different problem types, frameworks for instruction, and have designed, implemented, and tested multiple curricular components specifically designed to improve problem-solving ability [2–5].

When solving problems, it appears the defining distinction between expert and novice problem-solvers is found primarily in their metacognitive ability. Expert-like problem-solvers demonstrate an awareness of actual and potential flaws in their reasoning, they can articulate their reasoning process, and they have confidence to revise their thinking and approach when warranted [6]. The expert utilizes multiple resources towards solving a problem, whereas the novice is typically "stuck" in one type of framing [7].

Heller *et al.* have found that although most faculty believe metacognition to be primarily how students learn problem-solving, instructional goals and methods rarely match these beliefs [8]. This has led to the creation of new effective pedagogies designed to improve student problem-solving ability based on the theoretical framework of cognitive apprenticeship [9]. These pedagogies focus on modeling expert-like process, coaching and scaffolding of group problem-solving, and student articulation through written solutions. However, one of the key requirements for effective cognitive apprenticeship is reflection [11]. As a metacognitive process, reflection is a difficult activity to assign to students since it is ultimately a mental action that must be done by their own choosing [10].

In writing-intensive disciplines, such as rhetoric and composition, student peer-review is used to engage students in reflection on the writing and thinking of their peers, all in an effort to encourage metacognition [12]. Could peer-review of hand-written problem solutions in the physics class similarly improve student metacognition with respect to the problem-solving process? To start the process of answering this question, we have undertaken an investigation into the efficacy

of rubric-based peer-review and its potential as an assignable metacognitive task. Specifically, we have investigated the effect of peer-review on the development of students' problem-solving abilities: do students participating in peer-review become better problem-solvers compared to those that do not?

## II. THEORETICAL FRAMEWORK

Cognitive apprenticeship is the theoretical framework on which this study and the educational treatments described are based. In a traditional apprenticeship the apprentice learns processes through physical integration into the practices associated with the content area [13]. Cognitive apprenticeship borrows from traditional apprenticeship as an applied teaching technique for students constrained to the classroom [14].

The principal teaching methods of cognitive apprenticeship are summarized in Tab. I [15, 16]. Modeling, coaching, and scaffolding are the principal faculty-initiated methods for cognitive apprenticeship, which are designed to help students construct a conceptual model for content and develop a set of cognitive abilities through practice. Reflection and articulation serve to internalize the student's observations and experience, as well as aid in integrating new knowledge and problem-solving skills. Finally, exploration fosters independence and encourages autonomous problem formulations and solutions [17].

Effective pedagogies designed to improve student problem-solving ability in physics are based on the theoretical framework of cognitive apprenticeship [5, 9]. These pedagogies focus on modeling expert-like problem-solving process and then coaching and scaffolding student problem-solving through group sessions Heller *et al.* call Cooperative Group Problem-Solving (CGPS) [5]. Students engage in articulation through group discussion and via written solutions, and explore through individual challenge problems. In the CGPS pedagogical framework, a problem-solving rubric is typically used to both scaffolded expert-like problem-solving and serve as a reflective activity [18].

Similar to other groups, we use a CGPS rubric specific to the pedagogy to scaffold instruction in problem-solving pro-

TABLE I. The principle methods of cognitive apprenticeship [16].

Method	Description
Modeling	A subject expert explicitly demonstrates a task to the student.
Coaching	The expert observes the student attempting a task and gives them feedback and assistance at critical moments.
Scaffolding	Assistance is slowly withdrawn as the student gains new abilities
Reflection	The student reflects on their own performance in solving a problem through analysis and deconstruction.
Articulation	The student thinks about their own actions and explains them to others, making their knowledge explicit.
Exploration	Students investigate new methods and strategies by exploring new problems.

cess. In this study, we are attempting to determine if student peer-review of problems using the same rubric can serve as an assignable reflection task within the cognitive apprenticeship framework. The hypothesis is that increased reflection on process through use of the rubric in the assessment of peers will result in an increase in problem-solving process abilities.

### III. EXPERIMENT DESIGN

Figure 1 shows the study design, which consists of a pre/post-test quasi-experiment having non-equivalent groups. Students from a comprehensive university in the Midwest USA self-selected into a calculus-based mechanics introductory course with a total population of  $N = 45$ . The participants were a combination of engineering and science majors. All students attended the same lecture with the same instructor at the same time. Multimedia learning modules (MLMs) were used to "flip" the lecture component of the course, as described in detail elsewhere [19]. The participants completed the same pre-lecture MLMs and checkpoints with the same due dates.

Students self-selected into one of two CGPS discussion sessions that met on the same day one after the other. One session was assigned to be the treatment peer-review group (PR), and the other session was assigned to be the control non-peer-review group (Non-PR). Assignment of which section would be the control versus treatment group was random. CGPS sessions were facilitated by the same instructor and undergraduate learning assistant.

All students received the same modeling instruction and completed the same group problems during CGPS sessions and were assigned the same individual "challenge" problems due by the beginning of the next week's session. Challenge problems were to be completed individually, hand-written, digitized, and uploaded through a learning management sys-

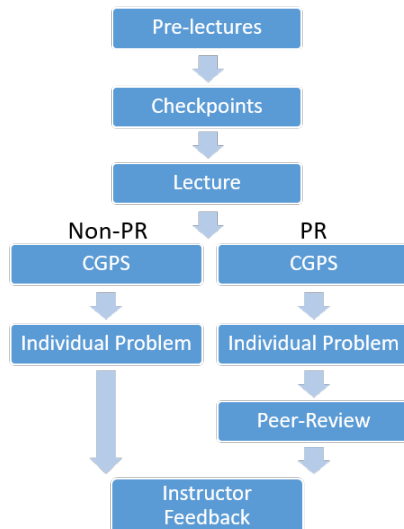


FIG. 1. Non-equivalent multiple-group design of the experiment. All students participated in the same pre-lecture and lecture activities. Students were split into two separate discussion sections for the CGPS activities.

TABLE II. FCI pre-score for the peer-review (PR) and non-peer-review (Non-PR) groups.

Group	N	Mean	SD	$p$
PR	20	0.574	0.237	0.482
Non-PR	24	0.524	0.227	

tem (LMS, Canvas). All challenge problems were graded by both the instructor and the learning assistant using a rubric we describe in detail elsewhere, with feedback posted to the LMS within one week [20]. The rubric was provided to all students on the first day of the course and its use was modeled during CGPS sessions. It consisted of 12 criteria across 3 factors: framing, planning and execution, and physics formalism, with its validity and reliability described in Reference [20]. Students in the PR group were additionally assigned the task of peer-review for three randomly selected classmates. Students used the same rubric to assess their peer's solutions.

Cognitive and preparation equivalence between the two groups was established using a two one-sided t-test (TOST) on the Force Concept Inventory (FCI) pre-scores. Table II shows the mean FCI pre-score and standard deviation (SD) for the PR group and the Non-PR. The assumption of equal variance was not violated ( $p > 0.05$ ), and the TOST shows 95% confidence ( $p < 0.05$ ) that the variance is within 10%.

Pre-study problem-solving process ability was measured by assigning and scoring a one-dimensional motion problem that needed no knowledge of specific physics formalism to successfully solve. Table III shows the mean and standard deviation (SD) for both the PR and Non-PR groups. The groups were not equivalent with respect to problem-solving ( $p <$

TABLE III. Problem-solving pre-score for the peer-review (PR) and non-peer-review (Non-PR) groups.

Group	N	Mean	SD	$p$
PR	18	29.3	4.80	0.019
Non-PR	22	32.2	2.54	

TABLE IV. One-way ANCOVA table with reliability-corrected pre-score as the covariate. Levene's test and normality checks were carried out and the assumptions met.

ANCOVA	F	$p$	$\eta^2$
Group	3.68	0.063	0.076
1D Motion Adjusted	7.94	0.008	0.163

0.05). Specifically, the PR group began the study with lower measured problem-solving process ability than the Non-PR group, which could result in a selection threat to the internal validity of the study, which we will discuss in more detail below.

The study was conducted over four weeks with challenge problems assigned weekly for topics in one-dimensional motion, two-dimensional motion, and force/acceleration. The post-study assessment of problem-solving process ability was conducted in the fourth week via assessment of a two-dimensional motion problem on the first exam.

#### IV. RESULTS

Figure 2 shows the average rubric scores on challenge problems for the PR and Non-PR groups over the course of the study, where problem numbers correspond to the following content topics: (1) one-dimensional motion, (2) two-dimensional motion, (3) force/acceleration, and (4) two-dimensional motion (exam). No significant change in rubric score was observed for the Non-PR group over the course of the study. A steady increase in average rubric score is observed over the four week period for the PR group.

A reliability-corrected one-way ANCOVA was conducted to compare the effectiveness of peer-review with pre-score (Challenge Problem 1) as the covariate. Table IV shows the ANCOVA table, where a statistically significant difference is seen between groups on mean post-test score controlling for pre-score ( $F = 7.94, p = 0.008$ ). Effect-size ( $\eta^2$ ) was found to be 0.163, which is considered a medium effect.

As discussed, the scoring rubric consisted of 12 criteria across 3 factors: framing, physics formalism, and planning and execution [20]. Figure 3 shows the average factor scores for the pre- and post-study problems. No significant change is observed in student problem framing for the PR group. A large increase in physics formalism scores is observed, while a small decrease in planning and execution scores are seen for

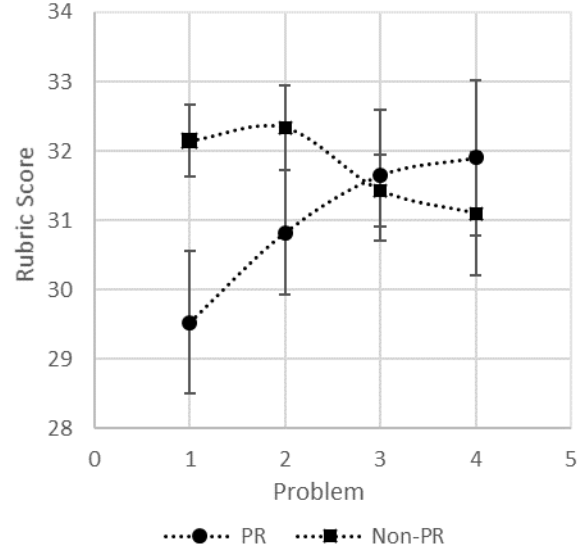


FIG. 2. Average rubric score for problems in (1) one-dimensional motion, (2) two-dimensional motions, (3) forces/acceleration, and (4) a two-dimensional motion exam question. Note: the zero has been suppressed so that the relatively small differences can be seen.

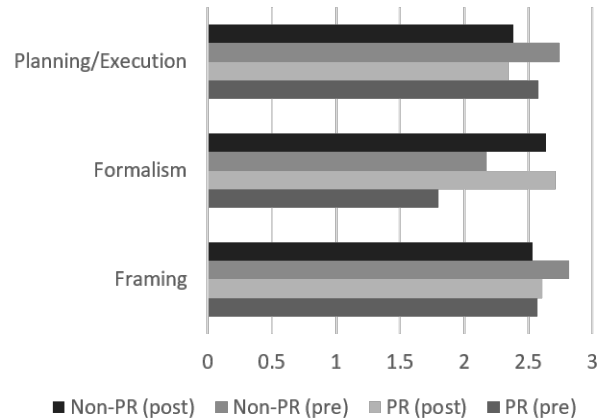


FIG. 3. Average factor score for Problem 1 (pre) and Problem 4 (post) for the PR and Non-PR groups.

both groups.

#### V. DISCUSSION

As seen in Fig. 2, the PR group began the study with lower measured problem-solving process rubric scores than the Non-PR group, but ended the study with higher rubric scores. However, the significance of the non-equivalence at the beginning of the study introduces the possibility of a selection threat, with the approach towards the Non-PR group suggesting a possible selection-regression or selection-

maturation threat to the internal validity of the study.

We do observe a steady increase in rubric score for the PR group, even when problems become more difficult, as suggested by Non-PR scores (see Problem 3, forces/acceleration in Fig. 2). The control group does not appear to change from pre- to post-assessment while the treatment group does. There is no evidence for selection-maturation due to the dormancy of the Non-PR group over the four week study. For selection-regression, we would not expect such a smooth approach with a final "cross-over," as seen in Fig. 2. Regression might explain why a low scoring treatment group approaches the control group's post-test score, but it does not explain the cross-over observed in Fig. 2. With that said, there was no significant difference when comparing Problem 4 scores ( $p > 0.05$ ). Therefore, we cannot completely eliminate the possibility of a selection-regression threat.

When looking at individual factors of problem-solving, we found the PR group's gain can be almost entirely contributed to improvements in physics formalism. The criteria for this factor includes defining the target quantity, explicitly stating known values and unknown values, physics-specific sketches (such as vector diagrams), and defining quantitative relationships. These criteria could be considered the most pedagogy-specific and the least problem-difficulty-dependent aspects of the rubric. These criteria may be less cognitively demanding, and therefore more susceptible to improvement via scaffolded practice and increases in content knowledge. The larger improvement achieved by the PR group compared to the Non-PR group could be attributed to increased time-on-task and mindful reflection on the rubric criteria.

Interestingly, there was no significant gain within the framing factor for the PR group, and a slight decrease in framing score for the Non-PR group. Proper problem framing is much more dependent on the integration of knowledge, practice,

and reasoning, which we could expect to be more difficult to develop in a short amount of time [7]. However, it should be noted that there is a large difference in problem difficulty between Problems 1 and 4. The PR group's slight gain in framing in comparison to the Non-PR group's decrease may be an indication that the peer-review process can improve difficulty-normalized problem framing. More study would be needed to provide support for this hypothesis.

## VI. CONCLUSION

At the beginning of this study, we asked the following research question: do students participating in peer-review become better problem-solvers compared to those that do not? A medium-sized positive effect ( $\eta^2=0.163$ ) on student problem-solving ability as measured by a problem-solving rubric was found over four weeks for students participating in rubric-based peer-review. This result could be explained by increased reflection-on-process, where the peer-review process acts as an assigned metacognitive task as outlined in our theoretical framework. The result could also be explained by increased time-on-task, since the PR group more than likely spent more time actively participating in the problem-solving process due to the inclusion of more assigned work compared to the Non-PR group. Furthermore, we cannot eliminate the possibility of a selection-regression threat to the internal validity due to the non-equivalence at the study outset.

We intend to continue this line of research with another similar study scheduled during a future course. Further research involving student interviews is also necessary to establish if the observed effect is the result of metacognitive processes. We are also interested in determining the relationship between metacognitive activity such as peer-review and problem-solving factors.

- 
- [1] National Academy of Sciences, *Rising Above the Gathering Storm* (National Academies Press, Washington DC, 2007).
  - [2] J.P. Mestre, *J. Appl. Dev. Psychol.* **23**, 9 (2002).
  - [3] P. Heller and M. Hollabaugh, *Am. J. Phys.* **60**, 637 (1992).
  - [4] K. VanLehn *et al.* *Int. J. Artif. Intell. Educ.* **15**, 147 (2005).
  - [5] P. Heller, R. Keith, and S. Anderson, *Am. J. Phys.* **60**, 627 (1992).
  - [6] A. Brown *et al.* in *Handbook of child psychology, Vol. 3 Cognitive development (4th ed.)*, edited by J. H. Flavell and E. M. Markman (Wiley, New York, 1983).
  - [7] K.M. Edmondson, and J.D. Novak, *J. Res. Sci. Teach.* **30**, 547-559 (1993).
  - [8] P. Heller *et al.* in *Proceedings of the Physics Education Research Conference, Rochester, NY 2001*
  - [9] J. Docktor and K. Heller, in *Proceedings of the Physics Education Research Conference, Ann Arbor, Michigan 2009*
  - [10] A.J. Mason, Ph.D. thesis, University of Pittsburgh, 2009.
  - [11] A. Collins, J. Brown, and A. Holum, *Am. Educator* **6** 38-46 (1991).
  - [12] K.B. Yancey, *Reflection in the Writing Classroom* (Utah State University Press, Salt Lake City, 1998).
  - [13] D. Pratt, *Five perspectives on teaching in adult and higher education* (Krieger Publishing Company, Malabar, FL, 1998).
  - [14] S. Barab and K. Hay, *J. Res. Sci. Teach.* **38**(1), 70-102 (2001).
  - [15] J. Brown, A. Collins, and P. Duguid, *Educational Researcher* **18**, 32 (1989).
  - [16] C. Moore, *Creating Scientists: Teaching and Assessing Science Practice for the NGSS* (Routledge, New York, 2017).
  - [17] D.B. Larkins *et al.* in *Proceeding of the 44th ACM Technical Symposium on Computer Science Education* (ACM, New York, 2013), p. 89.
  - [18] J. Docktor *et al.* *Phys. Rev. PER* **12**, 010130 (2016).
  - [19] J.C. Moore, *Educ. Sci.* **8**, 23 (2018).
  - [20] J.C. Moore, in *Proceedings of the Physics Education Research Conference, Washington, DC 2018*.