

Discussions on University Science Teaching: Proceedings of the Western Conference on Science Education

Volume 1 *Proceedings of the 2015 Western Conference on Science Education*

Article 9

2017

Enhancing Two-stage Collaborative Exams by Incorporating Immediate Feedback

Teresa (Terry) R. McCurdy
McMaster University, mccurdt@mcmaster.ca

Kim Volterman
McMaster University


Ralph Shiell
Trent University

Melec Zeadin
McMaster University

Kevin Dunn
McMaster University

See next page for additional authors

Follow this and additional works at: <http://ir.lib.uwo.ca/wcsedust>

 Part of the [Higher Education Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

McCurdy, Teresa (Terry) R.; Volterman, Kim; Shiell, Ralph; Zeadin, Melec; Dunn, Kevin; De Melo, Jason; and Helli, Peter (2017) "Enhancing Two-stage Collaborative Exams by Incorporating Immediate Feedback," *Discussions on University Science Teaching: Proceedings of the Western Conference on Science Education*: Vol. 1 , Article 9.
Available at: <http://ir.lib.uwo.ca/wcsedust/vol1/iss1/9>

This Article is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Discussions on University Science Teaching: Proceedings of the Western Conference on Science Education by an authorized editor of Scholarship@Western. For more information, please contact tadam@uwo.ca.

Enhancing Two-stage Collaborative Exams by Incorporating Immediate Feedback

Authors

Teresa (Terry) R. McCurdy, Kim Volterman, Ralph Shiell, Melec Zeadin, Kevin Dunn, Jason De Melo, and Peter Helli

Enhancing Two-stage Collaborative Exams by Incorporating Immediate Feedback

Terry McCurdy¹, Kim Volterman¹, Ralph Shiell², Melec Zeadin^{1,4}, Kevin Dunn³, Jason De Melo^{1,5}, Peter Helli¹

¹School of Nursing, McMaster University, Hamilton, Ontario, Canada

²Department of Physics and Astronomy, Trent University, Peterborough, Ontario, Canada

³Department of Chemical Engineering, McMaster University, Hamilton, Ontario, Canada

⁴McMaster Institute for Innovation & Excellence in Teaching & Learning, Hamilton, Ontario

⁵University Health Network, Toronto, Ontario, Canada

Abstract

Collaborative testing is becoming a widely used method of assessment in science education, as well as many other disciplines. While performance gains have been consistently reported, learning gains have not been quite as clearly documented. Learning gains are often measured by assessing the retention of tested concepts. We designed and conducted a study to assess whether retention was greater in a two-stage collaborative exam where students received immediate feedback rather than multiple-choice with no feedback during the collaborative stage. All students first wrote the multiple-choice test individually before re-writing the exam in small groups and receiving feedback via scratch cards based on either the first or second half of the test, with no feedback for the other half. Two weeks later, students were given a retention quiz with five questions concept-matched to each half of the original test. There was no significant difference in retention found between the randomized groups. Interestingly, however, we observed that those who received feedback on the first half of the collaborative stage of the exam performed better on the collaborative exam overall, despite there being no difference in individual marks between the two groups. This effect remained even when the test component with feedback was marked dichotomously, without part marks awarded for correctness on the second or third attempt.

Keywords: Collaborative testing, assessment, two-stage exam, immediate feedback

Introduction

Large class sizes are becoming the norm in post-secondary education, especially in introductory-level courses. Although group-work may be part of the teaching strategy, testing is more commonly conducted on an individual basis, and often utilizes a multiple-choice format with delayed marking via a Scantron® answer sheet. However, *collaboration during assessment*, when students are most prepared, may enhance learning and increase metacognitive skills (de Carvalho Filho, 2010), as well as offering additional benefits such as a reducing stress and anxiety (Pandey and Kapitanoff, 2011).

There has been much written about collaborative testing and it has taken many forms. Some of the literature is based on testing in groups only (Lusk & Conklin, 2003; Meseke, Nafziger, & Meseke, 2010; Zimbardo, Butler, & Wolfe, 2003), while some adopt a two-stage format, with individual testing first, followed by collaboration on a same, similar, or different test (Cortright, Collins, Rodenbaugh, & DiCarlo, 2003; Gilley & Clarkston, 2014; Leight, Saunders, Calkins, & Withers, 2012). These studies concluded that overall performance is enhanced by collaboration; there is currently insufficient evidence, however, to conclude that this necessarily translates into enhanced learning or retention. Through surveys and interviews,

students generally report a positive perception of this type of testing. One of the concerns of faculty when deciding to adopt collaborative testing is the increase in grades as a result of collaboration. The grade boost is more modest in the two-stage model, as the majority of the grade comes from the individual effort. When the group grade only comprises a small portion of the final grade, it is less likely that a student could pass a course with marks gained just from collaborating.

Evidence from the educational psychology literature also supports the two-stage model over collaborative testing alone, with better cognitive organization occurring when students first retrieve information individually before engaging collaboratively (Congleton & Rajaram, 2011, 2012). When students recall information only in a group, they can interfere with each other's retrieval strategies, resulting in "collaborative inhibition" (Congleton & Rajaram, 2011). This phenomenon is minimized by testing individually first. While collaborating, students not only benefit from talking about each question with others in their group, allowing for peer-led dispelling of misconceptions, they are also able to gain confidence and reaffirm their own knowledge when others express a similar understanding. Explaining concepts to others reinforces existing learning and arriving at a consensus strengthens interpersonal and group-work skills.

The benefits of immediate feedback have been separately documented (Epstein et al., 2002). Providing timely feedback can correct misconceptions and reduce perseveration of errors. With Immediate Feedback Assessment Technique (IF-AT) cards (Epstein, Epstein, & Brosvic, 2001) students remove a metallic coating for their selected option in a manner similar to scratching a lottery ticket, and with multiple attempts possible for each question, this allows an answer-until-correct approach. Students have a positive perception of the IF-AT cards because they provide multiple attempts to reassess their understanding and to possibly receive partial marks, and there is also an element of excitement in experiencing immediate confirmatory feedback. There is, however, a dearth of literature combining collaboration and immediate feedback, and meanwhile the ability for either collaborative testing (Ives, 2014; Leight et al., 2012) or IF-AT cards to enhance retention is controversial.

In order to explore this question of retention we designed a study, adapted from the two-stage model outlined by Gilley and Clarkston (2014), with the addition of immediate feedback in half of the collaborative stage of the test. In short, students first wrote the test individually prior to being randomized into groups of 3-5 in which they re-wrote the test collaboratively. For this study, the collaborative portion of the test was split; one set of groups (in room A) wrote the first half using Scantron® cards and the second half with IF-AT cards and the other set of groups (in room B) reversed the ordering (first half IF-AT, second half Scantron®) (see Figure 1).

We conducted the study in order to evaluate: (a) student retention of concepts, with and without immediate feedback in a collaborative setting, and (b) student perceptions of two-stage testing and the use of the IF-AT cards. We hypothesized that overall retention of course concepts would be superior with the use of immediate feedback and collaboration, rather than with collaboration alone. Furthermore, we predicted that students would have an overall positive perception (e.g. less stressful, more enjoyable, increased learning) of collaborative testing with immediate feedback, using the IF-AT cards.

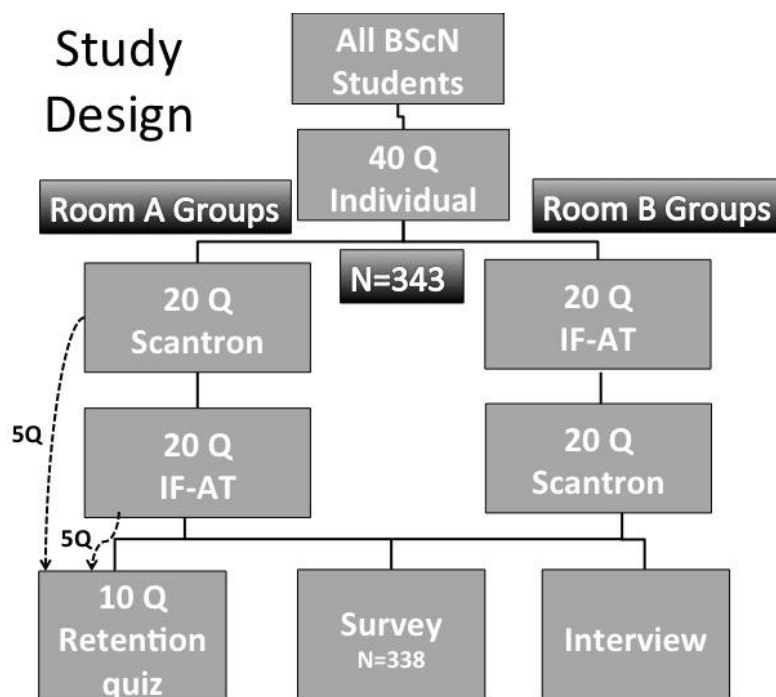


Figure 1. Flowchart of experimental design. The midterm consisted of 40 multiple-choice questions (MCQs) that were subsequently split into two 20-question halves for the collaborative portion. The retention quiz consisted of 5 concept-mapped questions from each 20-question half of the midterm, for a total of 10 MCQ questions; students in both room A and room B wrote the same retention quiz.

Methods

Participants

Participants were first year Bachelor of Science Nursing students enrolled in the Introductory Human Biochemistry course at one of three different academic sites. All participants provided written informed consent prior to study enrolment. Eighty-six percent of students from the course (343 out of 399) agreed to participate in the study, 95% of which were females.

Experimental Protocol

The midterm test used for this study protocol consisted of 40 multiple-choice questions (MCQs) and was part of the regularly scheduled evaluations for the Introductory Human Biochemistry course. Thus, the midterm was mandatory for all students. On the day of the test, students were separated into two rooms based on the first letter of their surname. Once the test began, students were first allotted 60 minutes to answer the 40 MCQs individually, using Scantron® cards (85% of grade). Following completion of the individual stage, students were allowed five minutes to form into their groups that had been randomly allocated prior to the test (4 groups of 3, 92 groups of 4; 2 groups of 5). Students were then allotted 40 minutes to answer the same 40 MCQs collaboratively in their groups, using a single IF-AT card per group (15% of each student's grade). Groups in Room A were instructed to answer the first 20 MCQs using Scantron® cards, and the last 20 MCQs using IF-AT cards. Groups in Room B executed this in reverse.

Approximately two weeks following the midterm, during their regularly scheduled tutorial sessions, students individually wrote a surprise retention quiz consisting of 10 MCQs. The first five MCQs were concept-matched with five MCQs from questions 1-20 on the midterm; the last five MCQs were similarly matched to questions 21-40. The retention quiz was administered as a pop quiz and, therefore, the students had no prior knowledge that the quiz would take place and so would not have been inclined to study. Marks on the retention quiz were not used for evaluation in the course; rather, one bonus mark was awarded for completion. Participants were also asked to voluntarily complete a survey comprising 10 questions using a five-point Likert scale to determine their perception of the collaborative midterm with and without IF-AT cards (N=338).

Statistical Analysis

All data were analyzed using SPSS Statistics Version 20.0. To determine differences in marks between room A and room B, independent t-tests were used. The significance level for all tests was set at $p < 0.05$. All data are presented as mean \pm SD, unless otherwise stated.

Results and Discussion

Collaborative Stage

One of the concerns of faculty when deciding whether or not to adopt collaborative testing is the potential for grade inflation as a result of collaboration. In the present study, we found that providing students with immediate feedback using the IF-AT cards resulted in modestly higher enhancements in grades compared with what we have previously observed in collaborative testing without immediate feedback (3.2% vs. 2.8% increase, respectively) (unpublished data). Although we did, in fact, see an increase in grades from the individual stage to the collaborative stage (71.9 ± 13.4 vs. 89.5 ± 7.1 , respectively), the overall grade boost is modest due to the fact that the majority (85%) of the grade comes from the individual stage of the test. As the midterm exam comprised 30% of the students' overall mark in the course, the actual increase in student grades with immediate feedback using IF-AT cards was negligible. Therefore, as the group grade only comprises a small portion of the final grade in the course, it is less likely that a student could pass a course with marks gained disproportionately from collaborating. Rather, this grade inflation might be justified if the students experience an increase in learning.

We did find an overall effect on midterm grades depending on whether students received immediate feedback on the first half or the second half of the collaborative stage of the midterm. Those groups who received immediate feedback during the first half of the collaborative stage scored better on the overall collaborative stage of the test compared to those groups who received immediate feedback on the second half ($90.5 \pm 6.8\%$ vs. $88.5 \pm 7.3\%$, respectively; $p = 0.02$). This effect remained when the IF-AT cards were re-marked dichotomously ($88.5 \pm 7.6\%$ vs. $86.3 \pm 8.4\%$, respectively, $p = 0.01$), removing part marks and providing a better comparison with the Scantron® part of the test. Since the groups were randomly allocated and there was no significant difference between room A and room B on the individual stage of the midterm ($p = 0.97$), there may have been several possibilities for why this might have occurred. First, the immediate feedback may have corrected any

misconceptions on the remainder of the collaborative stage of the test. Second, as similar concepts were addressed on both halves of the test, the immediate feedback may have bestowed greater self-confidence by confirming that students' thought processes were correct. Third, it is possible that immediate feedback may have revealed the most knowledgeable students in the group so that a greater weight was given to their input in the subsequent section without immediate feedback.

Retention Quiz

Whether students received immediate feedback on the first half or second half of the test did not significantly affect their performance on the first half ($p = 0.36$) or the second half ($p = 0.58$) of the retention quiz that was administered two weeks later in tutorial. Furthermore, although we had hypothesized that collaborative testing with immediate feedback would improve students' overall retention of course concepts compared to collaboration alone, overall performance on the retention quiz was not affected by which half of the midterm for which the student received immediate feedback. The average grades on the retention quiz were $54.8 \pm 19.3\%$ vs. $53.0 \pm 18.6\%$ ($p = 0.36$) for rooms A and B, respectively. Together, these findings suggest that immediate feedback had no impact on retention compared to collaboration alone. It is possible, however, that we were not able to detect an appreciable difference in retention due to either some limitations with the timing of the retention quiz or its format.

One shortcoming of the research design was that the date chosen for the midterm immediately preceded the spring reading break, a period when no classes are held. This limitation meant that retention had to be assessed over a longer time period than desired, but also that the students were away from school in the intervening time. Second, the bonus mark for writing the retention quiz was awarded regardless of performance, providing students with little or no motivation to do well. In the future, we would suggest conducting the retention quiz using the IF-AT cards as they are intrinsically motivating and provide a richer source of information. For example, the ability for students to answer again if incorrect on the first attempt can provide data on which answer was chosen first (if only two were chosen) and the analysis of part-marks may better inform the level of retention. The use of IF-AT cards would again have revealed correct answers, allowing another opportunity for learning and reinforcement from the immediate feedback. Finally, the number of quiz questions was a limitation in that it was difficult to assess a difference between groups. The questions were not assessed individually, however we do not believe that this further analysis would be likely to show a statistically significant difference since there was no significant difference in student performance between the first half and the second half of the retention quiz. This drawback was also exacerbated by the dichotomous nature of the marking and the lack of motivation to correctly answer the questions posed. Although a longer quiz may have allowed for question-specific analysis between the groups, a shorter quiz was deliberately chosen so as not to be onerous for students.

Qualitative Survey

The final aim of this study was to evaluate student perceptions of the two-stage testing process and the use of the IF-AT cards, for which we solicited student feedback immediately

following the retention quiz. The results indicated that students perceived they experienced enhanced learning through the combination of collaboration and immediate feedback. 83% of participants agreed or strongly agreed that the IF-AT cards would help them to retain more information than collaboration alone, while 92% felt that the use of IF-AT cards helped to correct their misunderstandings of course material. Importantly, 78% felt that their level of stress was reduced due to the ability to collaborate. This finding is in agreement with other data that has also suggested students experienced reduced stress and anxiety during collaborative testing (Breedlove, Burkett, & Winfield, 2004; Lusk & Conklin, 2003; Mitchell & Melton, 2003; Pandey & Kapitanoff, 2011). We found it encouraging that 70% of students stated they felt obliged to study in an effort to avoid letting down their other group members during the collaborative portion of the midterm. This increase in preparation prior to the midterm would have also benefitted students on the individual portion of the midterm exam. All survey responses had negligible relationship with individual and collaborative test marks, suggesting that students who did well (or poorly) on the midterm did not differ in their reactions to any of the testing techniques used here.

Conclusions

Although we were not able to confirm our original hypothesis that immediate feedback would improve retention above collaborative testing alone, it revealed some interesting benefits of immediate feedback in a collaborative setting. Superior performance on the collaborative stage of the test was conferred with the use of the IF-AT cards, if they were used first. Regardless of performance on the test, including the IF-AT cards in the testing helped decrease student stress during testing and helped to motivate students to study for the midterm.

Acknowledgements

The authors would like to thank the students for their participation. The study was approved by the Hamilton Integrated Research Ethics Board. Funding was provided by the McMaster Institute for Innovation & Excellence in Teaching & Learning (MIETL) and a President's Forward With Integrity Grant. The authors have no conflicts of interest to declare.

References

- Breedlove, W., Burkett, T., & Winfield, I. (2004). Collaborative testing and test performance. *Academic Exchange Quarterly*, 8(3), 36-41.
- Congleton, A. R., & Rajaram, S. (2011). The influence of learning methods on collaboration: prior repeated retrieval enhances retrieval organization, abolishes collaborative inhibition, and promotes post-collaborative memory. *Journal of Experimental Psychology: General*, 140(4), 535.
- Congleton, A., & Rajaram, S. (2012). The origin of the interaction between learning method and delay in the testing effect: The roles of processing and conceptual retrieval organization. *Memory & Cognition*, 40(4), 528-539.
- Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., & DiCarlo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Advances in Physiology Education*, 27(3), 102-108.

- Epstein, M. L., Epstein, B. B., & Brosvic, G. M. (2001). Immediate feedback during academic testing. *Psychological Reports, 88*(3), 889-894.
- Epstein, M. L., Lazarus, A. D., Calvano, T. B., & Matthews, K. A. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record, 52*(2), 187.
- de Carvalho Filho, M. K. (2010). Assessing changes in performance and monitoring processes in individual and collaborative tests according to students' metacognitive skills. *European Journal of Cognitive Psychology, 22*(7), 1107-1136.
- Gilley, B. H., & Clarkston, B. (2014). Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching, 43*(3), 83-91.
- Ives, J. (2014). Measuring the learning from two-stage collaborative group exams. *arXiv preprint arXiv:1407.6442*.
- Leight, H., Saunders, C., Calkins, R., & Withers, M. (2012). Collaborative testing improves performance but not content retention in a large-enrollment introductory biology class. *CBE-Life Sciences Education, 11*(4), 392-401.
- Lusk, M., & Conklin, L. (2003). Collaborative testing to promote learning. *Journal of Nursing Education, 42*(3), 121-124.
- Meseke, C. A., Nafziger, R., & Meseke, J. K. (2010). Student attitudes, satisfaction, and learning in a collaborative testing environment. *Journal of Chiropractic Education, 24*(1), 19-29.
- Mitchell, N., & Melton, S. (2003). Collaborative testing: An innovative approach to test taking. *Nurse Educator, 28*(2), 95-97.
- Pandey, C., & Kapitanoff, S. (2011). The influence of anxiety and quality of interaction on collaborative test performance. *Active Learning in Higher Education, 12*(3), 163-174.
- Zimbardo, P. G., Butler, L. D., & Wolfe, V. A. (2003). Cooperative college examinations: More gain, less pain when students share information and grades. *The Journal of Experimental Education, 71*(2), 101-125.