

Collaborative Testing: Evidence of Learning in a Controlled In-Class Study of Undergraduate Students

By Brett Hollis Gilley and Bridgette Clarkston

In collaborative two-stage exams, students complete a test as individuals and then immediately complete the same, or very similar, test in groups. We compared two-stage collaborative testing with individual testing to determine which format has a greater effect on student learning in an undergraduate Earth and Ocean Science course. A crossover design allowed students to participate in both the control (individual) and treatment (collaborative) conditions. In both the individual and collaborative conditions, students completed the same set of questions twice, which controlled for any potential performance gain caused by more frequent testing. Learning was measured as the change in students' individual performance on questions given in the individual stage and after the midterm, calculated as percent change and normalized change. When students were tested in groups, they showed significantly greater improvement on subsequent individual testing than when tested only as individuals. There was no significant difference in the amount of improvement experienced by "upper," "middle," or "lower" achieving students as categorized by their first-stage midterm score.

Most postsecondary institutions assess student learning with independent testing, that is, students complete the test on their own with no help from peers or outside resources. An alternative to this traditional format is the collaborative test, in which students work together in small groups to answer test questions. In the two-stage exam, perhaps the most common method of collaborative testing, students independently complete a test and then immediately complete the same, or similar, test again in groups of four; a proportion of each student's grade is assigned to the independent- and group-test sections (Cortright, Collins, Rodenbaugh, & DiCarlo, 2003; Stearns, 1996; Yuretich, Khan, Leckie, & Clement, 2001; see Leight, Saunders, Calkins, & Withers, 2012, and Zipp, 2007, for summaries of select collaborative testing literature). For recommendations on the use of two-stage exams in the classroom, see Appendix 1 (available at <http://www.nsta.org/college/connections.aspx>).

There are some common drawbacks to individual testing that can be diminished or even eliminated using collaborative testing. With a typical test, students often receive no feedback for days or even weeks while the test is graded; with a collaborative test, students receive immediate feedback from their peers while they write

the test. For instructors who encourage interaction in their classroom, solely individual testing (and treating student discussion of exam content as misconduct) conflicts with their teaching style, whereas a collaborative two-stage exam both assesses students individually and complements interactive teaching methods such as peer instruction, group-work activities, and other forms of interactive engagement (Lusk & Conklin, 2003). Collaborative testing can also reduce the anxiety (Lusk & Conklin, 2003; Zimbardo, Butler, & Wolfe, 2003) commonly associated with taking a test (Beilock, 2008; Hembree, 1988), increase positive relationships between students (Sandahl, 2010), improve student perception of the course (Shindler, 2004; Stearns, 1996), increase motivation to study (Shindler, 2004; Zimbardo et al., 2003), and decrease class dropout rates (Stearns, 1996).

The most commonly reported benefit of collaborative testing is improved test scores (e.g., Woody, Woody, & Bromley, 2008; Yuretich et al., 2001); however, several of these studies (e.g., Rao, Collins, & DiCarlo, 2002; Stearns, 1996) used the group's exam scores (rather than a posttest of individual students' performance) to indicate improvement—a method that fails to measure the impact of collaborative testing on individual student performance. Two studies used a subsequent posttest to measure the amount of

content retained by students following collaborative testing (Cortright et al., 2003, Leight et al., 2012). Cortright et al. (2003) reported significantly greater retention after 4 weeks on questions students had answered in groups compared with questions answered as individuals. Leight et al. (2012), in a study based on Cortright et al., reported no increase in retention when students were tested collaboratively. However, in both studies students completed the experimental questions three times in the collaborative condition and only twice in the individual condition,

which has the potential to confound their results with the well-established benefit of repeated testing (Karpicke & Roediger, 2008).

Our study was designed to build on the progress made by the previously mentioned studies, and our objectives were to (a) compare collaborative and individual testing where, in both conditions, students are exposed to the test questions the same number of times and given equal time to complete them; (b) use a low-stakes, follow-up assessment to remove the normalizing effect of student studying; and (c) measure

learning as the change in each student's individual performance on assessments given prior to and following the collaborative and individual tests.

Methods

Course context

The study was conducted during the 2012 summer section of EOSC 114—a nonmajors course about natural disasters offered by the Earth, Ocean, and Atmospheric Sciences Department at the University of British Columbia. The student population ($n = 98$) contained a mixture of junior

FIGURE 1

(A) flowchart of the experimental design used for both two-stage midterms. (B) An example of how the experimental design was experienced by a given student during each midterm (topic order reversed for one half of class).

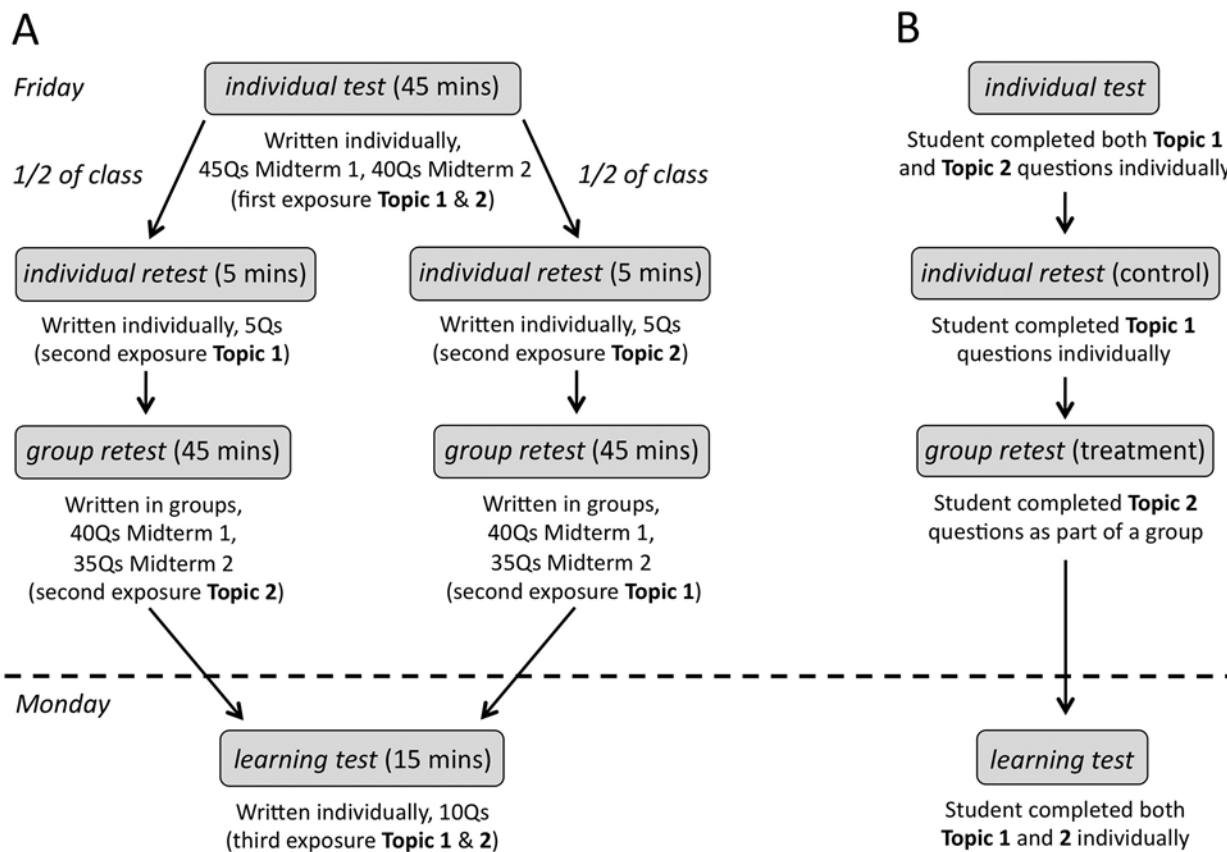


TABLE 1

**(A) Number and makeup of groups during each two-stage midterm.
(B) Number of students per topic included in the analyses because they completed both the midterm and follow-up learning test.**

	Topic seen by group	A			B
		Groups	Group makeup	Total students	Students included in analyses
Midterm 1	Topic 1	$n = 10$	8 groups of 4 2 groups of 5	42	38
	Topic 2	$n = 14$	14 groups of 4	56	41
Midterm 2	Topic 1	$n = 12$	1 group of 3 10 groups of 4 1 group of 5	48	35
	Topic 2	$n = 13$	2 groups of 3 11 groups of 4	50	36

(41% first and second year) and senior students (59% third year and above) from a wide variety of specializations from various parts of the university. The course spanned 3 weeks, with 2.5 hours of lecture each weekday and two midterms given during class on the first and second Friday. Though this is a nonstandard course structure, we believe it should not have a significant effect on the results of the study, as student learning is measured relative to other students taught in the same manor. The two-stage format was used for both midterm exams (each worth 30% of the students' final grade) and the final exam. For each midterm, 85% of the marks were allocated to the individual first stage and 15% to the collaborative second stage. A bonus of 1% was given to students as an incentive to complete the individual retest; no incentive was given to complete the learning test.

Experimental design

Our experiment took place during the course's two midterms, with each

midterm treated as a separate replicate. We used a quasi-experimental crossover design, meaning each student participated in both the control and the treatment condition, hereafter referred to as *individual* and *group*, respectively (Figure 1). The detailed experimental design was as follows: Students independently completed the first part of the midterm (hereafter called *individual test*; 45 questions for Midterm 1, 40 questions for Midterm 2). A subset of questions from the individual test were used for this study and divided into two topics based on content similarity (Topic 1 and Topic 2; five questions in each, 10 experimental questions total per midterm). The two Midterm 1 topics were "scales of disasters" and "earthquake waves"; the two Midterm 2 topics were "storm development" and "landslides." Topic 1 and Topic 2 scores from the individual test were used as the baseline assessment of students' knowledge. All experimental and nonexperimental questions were multiple choice. For representa-

tive questions, see Appendix 2 (available at <http://www.nsta.org/college/connections.aspx>).

Following the individual test, students rearranged themselves into groups of three to five. We then divided the class approximately into halves, each with a similar number of students and groups (Table 1).

Students then independently completed an individual retest of five questions (one half of the class received Topic 1 questions, the other half received Topic 2 questions). Repeat testing can cause people to remember material for longer—a phenomenon known as the "testing effect" (Butler & Roediger, 2007; Karpicke & Roediger, 2008; Roediger & Karpicke, 2006); the individual retest ensured that students were tested an equal number of times in the individual and group condition. Thus, any learning gain observed would not be due to answering the same question twice during a test (inherent to any two-stage test), but instead be the result of students working collaboratively.

Students then worked in groups to complete the group retest, which was identical to the individual test except that it lacked the questions (either Topic 1 or Topic 2) that the students' saw during the individual retest; that is, it contained five fewer questions than in the individual test (40 questions for Midterm 1, 35 questions for Midterm 2). During the group retest, each group was given a single paper copy of the test to complete together, and all members received the same score.

A learning test administered 3 days later as a surprise "pop quiz" contained all Topic 1 and Topic 2 questions (10 questions total) and was used as the follow-up assessment of students' knowledge. Students were not informed of the follow-up learning test until it was administered. The instruc-

tor gave students neither feedback nor the correct answer to any midterm questions until after the learning test.

Data analysis

We used two primary metrics to assess the impact of collaborative testing on learning, each of which is defined in detail below: percentage learning gain and normalized change. All statistical analyses were performed with JMP 9.0.2 (SAS Institute Inc., Cary, NC).

The data set for each midterm was analyzed separately and included only the scores from students who completed both the individual test and the learning test (Midterm 1: $n = 79$; Midterm 2: $n = 71$). Each student's individual test score was paired with his or her learning test score for each of the Topic 1 and Topic 2 question sets prior to any analyses. The class data set was then divided according to whether students had answered questions from a given topic during the individual retest (i.e., control) or group retest (i.e., treatment).

We calculated percentage learning gain for each student as the difference in their performance on the experimental questions during the baseline individual test and the follow-up learning test:

$$[\% \text{ correct}_{\text{learning test}} - \% \text{ correct}_{\text{individual test}}]$$

To compare the group and individual conditions, we calculated the difference in the amount of change that occurred within each condition for each student:

$$\frac{[\% \text{ correct}_{\text{learning test}} - \% \text{ correct}_{\text{individual test}}]_{\text{group}} - [\% \text{ correct}_{\text{learning test}} - \% \text{ correct}_{\text{individual test}}]_{\text{individual}}}{[\% \text{ correct}_{\text{learning test}} - \% \text{ correct}_{\text{individual test}}]_{\text{individual}}}$$

The effect size between the mean class scores for the group and indi-

vidual conditions was calculated for each midterm using Cohen's d equation and a pooled standard deviation (Coe, 2002). The uncertainty around each effect size was calculated as standard deviation using the equation:

$$SD = \sqrt{\frac{N_{\text{group}} + N_{\text{individual}}}{N_{\text{group}} \times N_{\text{individual}}} + \frac{d^2}{2(N_{\text{group}} + N_{\text{individual}})}}$$

Our second measure was normalized change. Normalized change, c , is a measure of a student's gain or loss relative to their maximum potential gain or loss (Marx & Cummings, 2007). When a student's performance improved from the individual test to the learning test, the normalized change for that student was calculated as follows: $c = 100 * ([\text{learning test} - \text{individual test}] / [100 - \text{individual test}])$. When a student's performance declined, his or her normalized change was calculated as $100 * ([\text{learning test} - \text{individual test}] / \text{individual test})$. If a student's individual test and learning test scores were equal, normalized change was recorded as 0. For each of the individual and group conditions, the individual c -scores were used to calculate the average normalized change for the class, c_{ave} . Standard errors for the reported c_{ave} values were calculated to provide a coarse depiction of the spread of values (Marx & Cummings, 2007). Students who earned 100% on both the individual test and learning test assessments were removed from the normalized change dataset because the performance of these students was considered to be beyond the scope of the assessments used in this study (Marx & Cummings, 2007); the final normalized change dataset included 67 students for Midterm 1 and 53 students for Midterm 2.

Results

The percentage learning gains results are summarized in Table 2. Unless otherwise stated, uncertainties are reported as standard error of the mean. We first compared the percent scores for the separate Topic 1 and Topic 2 question sets, for each midterm. One question from Topic 2 on Midterm 2 was dropped post hoc from the data set (and from the midterm grading) after it was determined that the question could be interpreted in such a way that more than one of the answer choices was correct. Student performance was similar for the two topics on Midterm 1 (Topic 1: $64.1\% \pm 2.1$, Topic 2: $65.3\% \pm 2.0$; paired t -test, $t = 0.46$, $df = 78$, $p = .65$) but not on Midterm 2 (Topic 1: $69.9\% \pm 2.5$, Topic 2: $55.3\% \pm 2.7$; paired t -test, $t = 4.56$, $df = 70$, $p < .001$). However, the cross-over design ensured that each topic was part of both the group and individual condition so the differences should not influence the interpretation of the results. We pooled the scores for each topic and divided them into two categories: questions answered in groups during the group retest (i.e., group condition) and questions answered during the individual retest (i.e., individual condition; Table 2).

Both midterms show the same pattern—a statistically significant improvement in student performance between the individual retest and the learning test in the group condition (Midterm 1: $13.1\% \pm 2.1$; Midterm 2: 13.2 ± 2.8 ; Table 2) and a statistically insignificant improvement in the individual condition (Midterm 1: $3.8\% \pm 1.5$; Midterm 2: $4.0\% \pm 2.1$; Table 2). The improvement in individual student scores on the learning test indicates that (a) during the collaborative group retest, students acquired correct knowledge that they did not have previously, and (b) learning was sufficient enough

for students to remember it 3 days later on an individually written surprise quiz.

For the learning due to the collaborative stage of the midterm, the observed effect sizes of 0.54 (± 0.16 SD) and 0.39 (± 0.17 SD) for Midterm 1 and 2, respectively, are notable considering that the type of intervention used (group component on a midterm) is of short duration and easily incorporated into a regular course. The effect size of 0.54 means that the score of the average student in the group condition is 0.54 standard deviations above that of the average student in the individual condition for Midterm 1 (Coe, 2002).

Learning gain was also calculated as the mean normalized change for the class, c_{ave} , which for Midterm 1 was $38.9\% \pm 4.7$ (group) and $12.3\% \pm 3.7$ (individual) and for Midterm 2 was $34.6\% \pm 7.1$ (group) and $13.9\% \pm 5.1$ (individual; Figure 2). An individual c -score of 33.0 in this study was equiva-

lent to a student's score improving by one question out of a possible improvement of three questions (the lowest positive c -score possible was 25.0, a one-out-of-four question improvement). The class c_{ave} values for each midterm indicate that, on average, when tested collaboratively students learned the correct answer to over one third of the questions that they had previously answered incorrectly as individuals (and had retained this learning 3 days following the midterm).

We provide a more detailed breakdown of the results from Midterm 1 in Figures 3 and 4. The results of Midterm 2 follow the same pattern, but they are not shown for the sake of brevity. The mean score on the midterm's individual test was almost identical for both the individual and group conditions, and only when tested collaboratively was there a significant improvement in student performance on the follow-up

learning test (Table 2). Interestingly, the scores on the individual learning test are similar to what the groups achieved on the same questions during the midterm's group retest (independent t -test, Midterm 1: $t = 0.41$, $df = 38$, $p = .68$; Midterm 2: $t = 0.35$, $df = 50$, $p = .73$; Figure 3). The similarity between group performance and subsequent individual performance suggests that nearly all of the improvement in midterm scores from the individual test to the collaborative group retest resulted in learning by the individual students that was subsequently measured on the learning test. The same pattern was observed when the data were broken down by question (Figure 4).

We also investigated whether two-stage exams were more beneficial for different groups of students based on their performance on the individual test of Midterm 1. We divided the students in the normalized change data set into

TABLE 2

Class performance (mean % \pm standard error) during each stage of the midterms.

Condition	Individual test (IT)	Retest (individual or group)	Learning test (LT)	Difference LT – IT	p -value (paired t -test)
Midterm 1					
Group	64.6 \pm 1.9	79.2 \pm 3.1	77.7 \pm 1.7	13.1 \pm 2.1	$t = 6.27$, $df = 78$, $p < .0001$
Individual	64.8 \pm 2.1	66.3 \pm 2.1	68.6 \pm 2.1	3.8 \pm 1.5	$t = 2.55$, $df = 78$, $p = .0127$
Difference between conditions				9.3 \pm 2.6	$t = 3.64$, $df = 78$, $p = .0005$
Effect size				0.54 \pm 0.16 (SD)	—
n				79	
Midterm 2					
Group	62.5 \pm 2.7	77.4 \pm 3.9	75.7 \pm 2.8	13.2 \pm 2.8	$t = 4.67$, $df = 70$, $p < .0001$
Individual	62.6 \pm 2.7	64.4 \pm 2.3	66.6 \pm 2.7	4.0 \pm 2.1	$t = 1.94$, $df = 70$, $p = .0561$
Difference between conditions				9.2 \pm 3.6	$t = 2.54$, $df = 70$, $p = .0132$
Effect size				0.39 \pm 0.17 (SD)	—
n				71	

Note: The differences in student performance between the baseline individual test and follow-up learning test assessments were compared within and between the group and individual conditions (p values). Midterm scores are shown for the relevant experimental questions only; scores for the full midterms were: 70.5% \pm 1.3 for Midterm 1 individual test, 88.4% \pm 1.3 for the group retest, and 73.1% \pm 1.2 for Midterm 2 individual test, 77.5% \pm 1.1 for the group retest.

three quantiles of roughly equal size and based on students' scores on all questions during the individual test: lower ($\leq 50\%$; $n = 20$), middle (60% ; $n = 22$), and upper ($\geq 70\%$; $n = 25$), respectively. We then compared the difference in normalized change between the group and individual conditions within each quantile, as well as the difference in normalized change between quantiles for each condition (Figure 5). Two-stage exams were more beneficial than individual testing for all students, regardless of prior performance on the individual test of the midterm. A two-way analysis of variance (ANOVA) revealed no interaction between condition (group and independent) and quantile ($F_{(5, 128)} = 0.15, p = .86$) on average normalized change (c_{ave}) in the group condition, which suggests that the lower, middle, and upper students benefited equally from two-stage exams in this study.

Discussion

The mean scores on the collaborative group retest of the exam are the highest of those on any assessment in this study. However, we have also demonstrated that individual student knowledge increases substantially during a two-stage exam. Most of the improvement seen between the individual test and the group retest is retained by students, as reflected in their individually written learning test. With no feedback other than that provided by group discussions with peers and no specific additional reason to study in the interim, most students correctly answered nearly half the questions they had answered incorrectly on the individual test if they were part of the collaborative exam. The fact that students' scores on the learning test followed closely what the groups scored suggests that the learning was on a deeper level than simple rote memorization. These results differ from pre-

FIGURE 2

The improvement in student performance was greater for the group condition compared to the individual, measured as average normalized change, c_{ave} , for the class between the baseline (individual test) and follow-up (learning test) assessments. Normalized change was calculated for each student before determining the mean of the class. Each midterm was analyzed separately. Bars represent standard error.

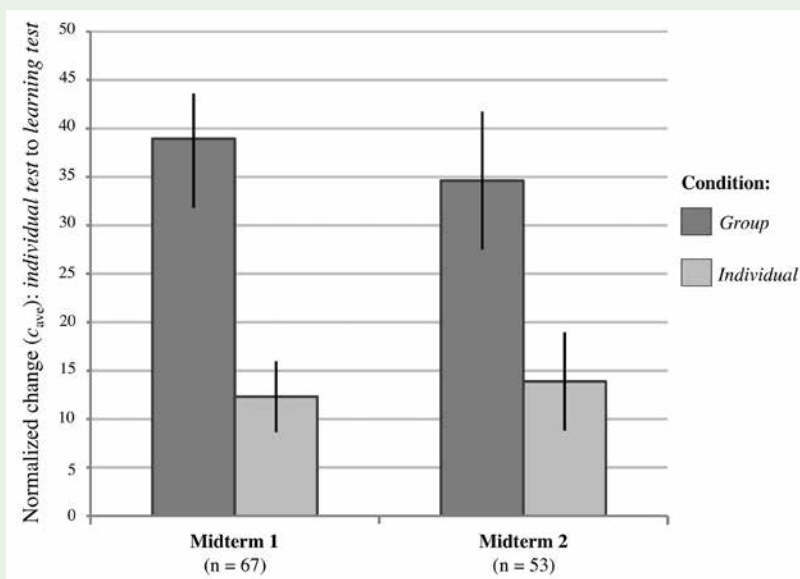
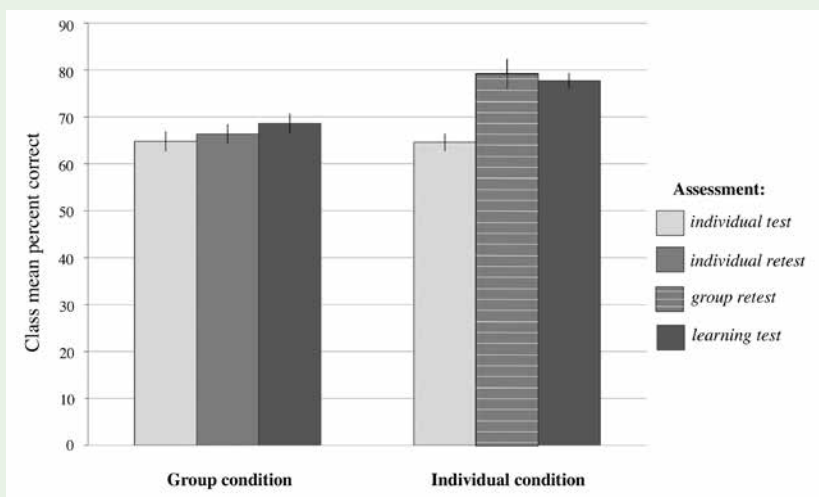


FIGURE 3

The improvement in student performance on the follow-up learning test was similar to (i.e., not significantly different from) what the groups achieved on the same questions during the group retest. Bars represent standard error.



vious studies such as Cortright et al. (2003) and Leight et al. (2012), because students in the individual and collaborative conditions saw questions an equal number of times and because there was a shorter time period between the two-stage exam and the learning test, 3 days in this study versus 4 weeks for Cortright et al. and several weeks, the time between midterms, for Leight et al.

Two-stage exams are not simply a process of the top students providing answers to the rest of the group; most groups score higher than any individual in the group. There are substantial and similar normalized learning gains for all groups of students as a result of the collaboration; indeed, in this study only four individuals scored equal to or higher than their group over both midterms. This finding addresses a concern commonly voiced about

collaborative testing (e.g., Leight et al., 2012)—that they benefit only the lower students. This is not the case based on our evidence.

Conclusion

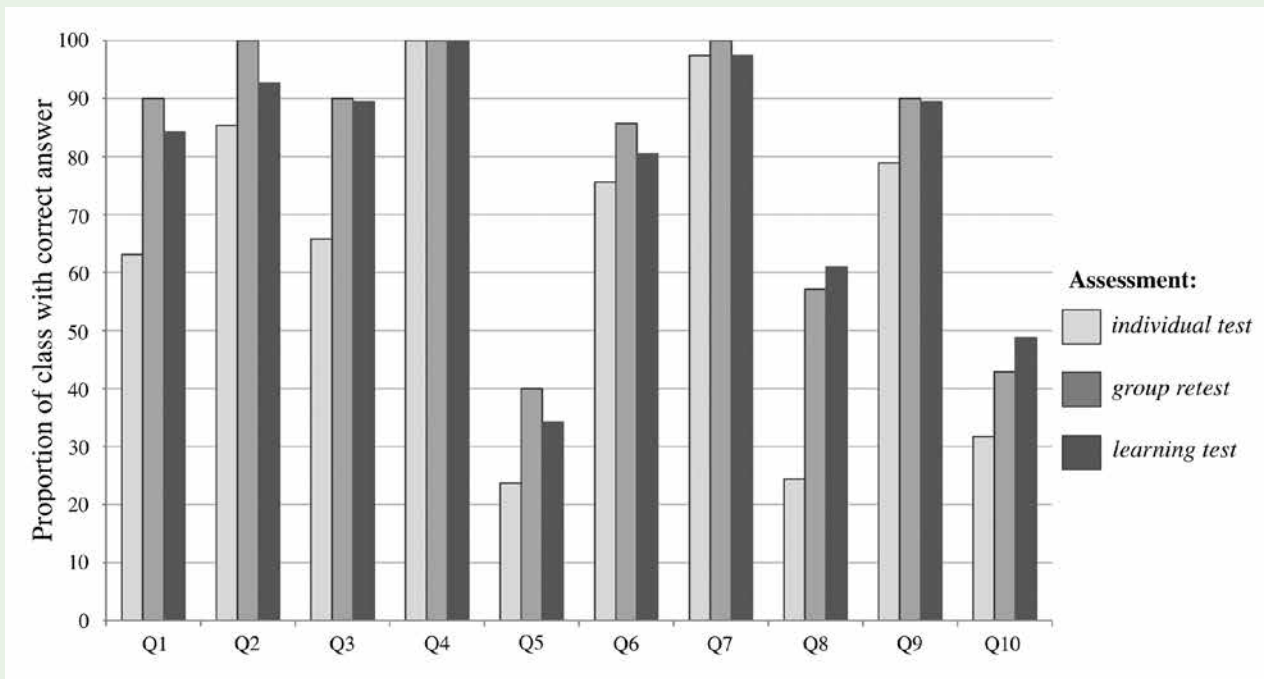
Two-stage exams serve as powerful learning experiences in addition to providing the traditional summative assessment of a traditional exam. They exploit what can be a highly productive learning environment, including high student motivation, high incentives for collaboration and communication, and immediate feedback on individual performance. Collaborative testing is not a new idea (e.g., Heller & Hollabaugh, 1992; Lusk & Conklin, 2003) and has been used in many disciplines, including geosciences (Yuretich et al., 2001), nursing (Sandahl, 2010), biology (Leight et al., 2012), and psychology

(Drouin, 2010). Though still limited in use (e.g., Wiggs, 2011), collaborative testing is slowly becoming more common at some institutions. For example, two-stage exams are now used at the University of British Columbia in over 20 courses across physics; chemistry; biology; math; statistics; Earth, ocean, and atmospheric sciences; computer science; forestry; and land and food systems. They occur in a variety of class formats as well, from 450-student, first-year lectures to graduate-level classes with fewer than five students.

The implementation of two-stage exams is relatively simple; they can easily be inserted into a wide variety of classes with very little disruption (recommendations for the use of two-stage exams can be found in the appendix at www.nsta.org/college/connections.aspx). The exam structure

FIGURE 4

The mean student performance during each assessment of the group condition, broken down by experimental question ($n = 38$ for Questions 1, 3, 5, 7, 9, 10 and $n = 41$ for Questions 2, 4, 6, 8, 10).



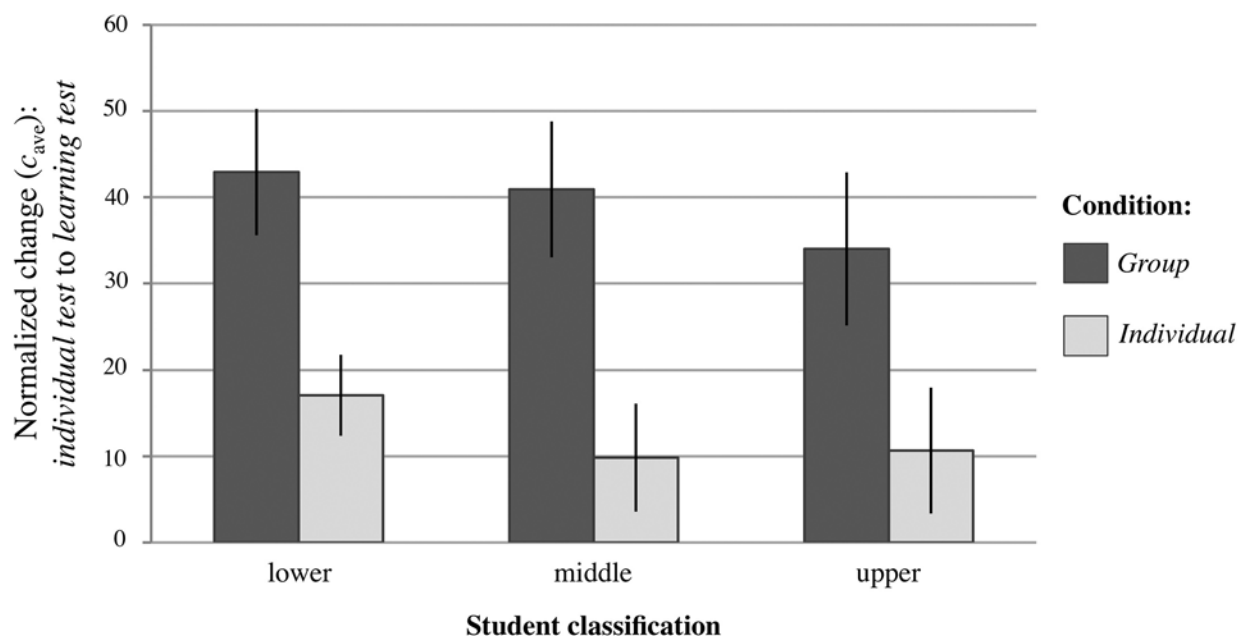
itself follows traditional formats; the only change required is shorter length to allow time for group discussion. For a relatively simple and short intervention, this style of exam shows a large benefit; the increase in student learning in a two-stage exam more than makes up for any small loss of evaluative material. Because two-stage exams are relatively easy to apply and can benefit both instructors and students alike, we believe they should be in wide use across postsecondary education. ■

References

- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*, 17, 339–343.
- Butler, A. C., & Roediger, H. L., III. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19, 514–527.
- Coe, R. (2002, September). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at the annual conference of the British Educational Research Association, University of Exeter, England.
- Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., & DiCarlo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Advanced Physiological Education*, 27(3), 102–108.
- Drouin, M. A. (2010). Group-based formative summative assessment relates to improved student performance and satisfaction. *Teaching of Psychology*, 37(2), 114–118.
- Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics*, 60, 637–644.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58, 47–77.
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968.
- 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 Science Education*, 11, 392–401.
- Lusk, M., & Conklin, L. (2003). Collaborative testing to promote learning. *The Journal of Nursing*

FIGURE 5

Normalized change (c_{ave}) for three classifications of students (lower, middle, and upper) based on their baseline individual test scores. Bars represent standard error.



- Education*, 42(3), 121–124.
- Marx, J. D., & Cummings, K. (2007). Normalized change. *American Journal of Physics*, 75, 87–91.
- Rao, S. P., Collins, H. L., & Dicarolo, S. E. (2002). Collaborative testing enhances student learning. *Advances in Physiology Education*, 26, 37–41.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- Sandahl, S. S. (2010). Collaborative testing as a learning strategy in nursing education. *Nursing Education Perspectives*, 31(3), 142–147.
- Shindler, J. V. (2004). “Greater than the sum of the parts?” Examining the soundness of collaborative exams in teacher education courses. *Innovative Higher Education*, 28, 273–283.
- Stearns, S. A. (1996). Collaborative exams as learning tools. *College Teaching*, 44(3), 111–112.
- Wiggs, C. M. (2011). Collaborative testing: Assessing teamwork and critical thinking behaviors in baccalaureate nursing students. *Nurse Education Today*, 31, 279–282.
- Woody, W. D., Woody, L. K., & Bromley, S. (2008). Anticipated group versus individual examinations: A classroom comparison. *Teaching of Psychology*, 35, 13–17.
- Yuretich, R. F., Khan, S. A., Leckie, R. M., & Clement, J. J. (2001). Active-learning methods to improve student performance and scientific interest in a large introductory oceanography course. *Journal of Geoscience Education*, 49(2), 111–119.
- 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.
- Zipp, J. F. (2007). Learning by exams: The impact of two-stage cooperative tests. *Teaching Sociology*, 35, 62–76.

Brett Hollis Gilley (bgilley@eos.ubc.ca) is a CWSEI teaching and learning fellow in the Department of Earth Ocean and Atmospheric Science and **Bridgette Clarkston** is a science teaching and learning fellow in the Department of Life Sciences, both at the University of British Columbia, Vancouver, British Columbia, Canada.

SHARE YOUR IDEAS!

Have an idea for an inspiring presentation or workshop on science education?
Submit a session proposal today for...

NSTA's 2014+2015 CONFERENCES

to enthuse and stimulate our
community of educators!

2014

Conferences on Science Education

- Richmond, Virginia**
October 16–18, 2014
- Orlando, Florida**
November 6–8, 2014
- Long Beach, California**
—in Collaboration with CSTA
December 4–6, 2014

Proposal Deadline:
1/15/2014

2015

National Conference on Science Education

- Chicago, Illinois**
March 26–29, 2015

Proposal Deadline:
4/15/2014

For more information, visit
www.nsta.org/conferences