# Does collaborative testing increase students' retention of concepts?

Brett Gilley (bgilley@eos.ubc.ca) and Dr. Bridgette Clarkston (bclarkst@zoology.ubc.ca)
Carl Wieman Science Education Initiative, University of British Columbia, Vancouver Canada

## Background

Learning through collaboration, even in a testing situation, has many benefits stemming from peer-to-peer interactions. A collaborative test, hereafter called a two stage exam, typically has the following format (Stearns 1996, ):

**1st Stage:** Students write exam as individuals.

**2nd Stage:** Groups of 3–5 students immediately complete a second identical (or very similar) exam. The 2nd Stage typically takes much less time.

Two-stage exams are reported to improve retention of concepts by individual students (Cortright *et al.* 2003) in addition participants report reduction in test anxiety (Russo and Warren 1999), greater motivation to study and think critically during a two stage exam (Shindler 2004). No previous study has tested for retention while controlling for the additional "time-on-task" of a two-stage exam format, in which students are exposed to the same questions twice.

## Research Questions

1) Does collaboration during a two-stage exam increase students' retention of concepts more than a test written individually?

2) What, if any, specific effects does collaboration during a test have on students' retention of concepts?

## Methods

*The course*

— Earth and Ocean Sciences 114: Natural Disasters (non-majors)
— Three week summer course (2.5 hrs classes, 5 days / week)
— 98 students, 59 % 1st- and 2nd-year, 41% 3rd-year and above.
— Study occurred over two midterms, each held on a Friday, with the Retention Test the following Monday.

*Experimental Set-up: A Cross-Over Design*

— Midterms were two-stage exams as described above, with two extra parts:
- **Individual Study**. Students re-do, as individuals, five 1st stage questions. Acted as the control treatment. Used to make sure students in the individual mode work on questions for the same amount of time (Figure 1).
- **Retention Test.** Individually-written quiz of 10 questions. Used to measure students' retention of concepts (Figure 1).

— Within each midterm, the crossover design was as follows:
- 10 experimental questions, in two topics by content similarity (Topic 1, Topic 2).
- 1st stage of midterm. All students complete the same test (45 Qs, Topic 1=Q1, 3, 5, 7, 9, Topic 2=Q2, 4, 6, 8, 10).
- Individual Study. Half of the class had Topic 1, the other Topic 2.
- 2nd Stage of the midterm. Groups of 3–5 re-wrote the 1st stage, *omitting whichever Topic the group's members saw during Individual Study* (40 Qs).
- Three days following the midterm, students completed the Retention Test assessment of 10 questions, identical to Topic 1 and Topic 2 questions.

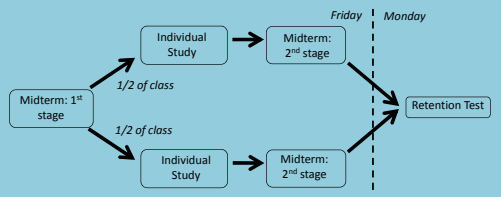


Figure 1: Flowchart of experimental design (see Methods for details)

## Results and Discussion

1) Does collaboration during a two-stage exam increase student's retention of concepts?

*For both midterms, the Collaborative mode resulted in significantly greater retention of concepts by students compared to the Individual test mode.*

Table 1 (Below): Students' Retention Test scores were significantly higher than their 1st stage scores (i.e., prior to any testing intervention) only when tested using the Collaborative mode (blue boxes). This increase in student retention was significantly greater for the Collaborative mode vs. the Individual mode (orange boxes). # of questions per column is 5. An * denotes significance.

Midterm 1 (n = 79)

| Test Mode | 1st Stage | Retention Test | Difference | *p* value |
|---|---|---|---|---|
| Individual (I) | 64.8 ± 2.1 | 68.6 ± 2.1 | 3.8 ± 1.5 | 0.1707 |
| Collaborative (C) | 64.6 ± 1.9 | 77.7 ± 1.7 | 13.1 ± 2.1 | 0.001* |
| Diff C – Diff I | 9.3 ± 2.6 | *p* < 0.001* | | |

Midterm 2 (n = 71)

| Test Mode | 1st Stage | Retention Test | Difference | *p* value |
|---|---|---|---|---|
| Individual (I) | 62.6 ± 2.7 | 66.6 ± 2.7 | 4.0 ± 2.1 | 0.251 |
| Collaborative (C) | 62.5 ± 2.7 | 75.7 ± 2.8 | 13.2 ± 2.8 | 0.0014* |
| Diff C – Diff I | 9.2 ± 3.6 | *p* = 0.0137* | | |

*The mean normalized learning gain was greater when using a Collaborative (C; purple bars) vs. Individual (I; blue bars) test mode, for both midterms*

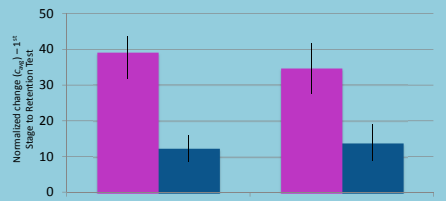


Figure 2 (Right): Graph shows the normalized change ($C_{avg}$) between the 1st stage of the midterm and the follow-up retention test for each test mode. Midterm 1 (C: 38.9±4.7 SEM, I: 12.3±3.7 SEM) and Midterm 2 (C: 34.6±7.1 SEM, I: 13.9±5.1 SEM). Normalized change is a measure of each student's gain in test score relative to that individual's maximum potential gain.

## Results and Discussion *continued*

2) What, if any, specific effects does collaborating during a test have on students' retention of concepts?

*The potential gain in retention for each student may be limited by their group's score*

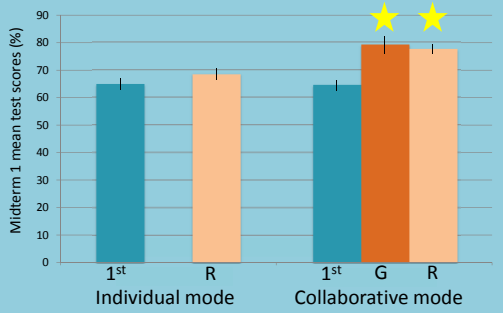


Figure 3 (above): For the collaborative mode, the class mean score (%) on the Retention Test was similar to the class mean score when working in groups during the 2nd stage of the midterm (yellow stars). Bars represent the class's mean score (%±SEM), for each midterm section, on questions relevant to the Individual and Collaborative modes. 1st = individuals' score during 1st stage of midterm, G = score achieved by groups during 2nd stage, R = individuals' score on retention test. Only midterm 1 data shown, but midterm 2 showed same general result.

*Comparing each question separately, students' retention test scores still appear to be influenced by their group's score*

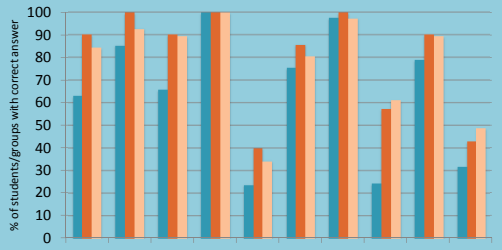


Figure 4 (Above): The class's mean score (%) during each midterm section, broken down by each question and given for the collaborative mode only. Blue = 1st stage, orange = 2nd stage, light orange = Retention Test. Only midterm 1 data shown, but midterm 2 showed same general result.

## Results and Discussion *continued*

*When comparing normalized gain by quantiles of the class (based on midterm mark) collaborative testing benefits all students equally, regardless of pre-intervention test performance*

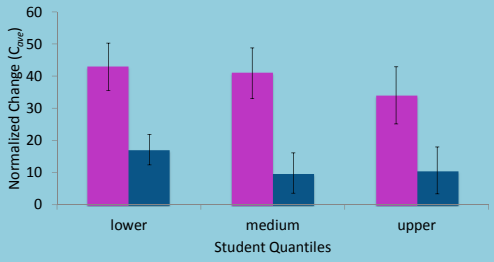


Figure 5 (above): Students in the lower, medium and upper performance quantiles achieved a similar gain in retention during the collaborative test mode. Bars represent the normalized change (i.e., retention gain between the 1st stage and Retention Test) for students during the collaborative (purple) and individual (blue) test modes. Students were divided into three quantiles ("lower", "medium" and "upper") based on their 1st stage scores. [A two-way ANOVA test was also used to determine a non-significant effect ($p$=0.104) of quantile on students' Retention Test scores.] Only midterm 1 data shown, but midterm 2 showed same general result.

## Conclusions

- Students showed a significantly higher gain in retention when tested in a collaborative setting over a traditional, individual-written test setting.
- Students' retention appears to be influenced by the performance of their group in the 2nd stage of the exam.
- Regardless of their performance prior to the two stage exam, all students appear to benefit equally when tested collaboratively.

## References

Cortright, R.N., Collins, H.L, Rodenbaugh, D.W., and DiCarlo, S.E. (2003) **Student retention of course content is improved by collaborative-group testing** Physiol. Edu. 27: 102-108 pp, 2003
Russo, A. and Warren, S.H. (1999) **Collaborative Test Taking** College Teaching, Vol. 47, No. 1 (Winter, 1999), pp. 18-20
Shindler, J.V. (2004) **"Greater Than the Sum of the Parts?" Examining the Soundness of Collaborative Exams in Teacher Education Courses** Innovative Higher Education Volume 28, Number 4

## Acknowledgments

The authors would like to thank the Carl Wieman Science Education Initiative for funding and support for this project. We would also like to thank Dr. Carl Wieman, Francis Jones, Dr. Ido Roll, Dr. Louis Deslaurier, Dr. Mandy Banet, Dr. Laura Wier, Dr. James Day, Dr. Georg Reiger, Dr. Cynthia Heiner, Dr. Sarah Gilbert, Dr. Warren Coad, all the current and former Science Teaching and Learning Fellows in the Carl Wieman Science Education Initiative and most importantly Students in the 2012 Summer Section of EOSC 114