Peer Instruction and Lecture Tutorials Equally Improve Student Learning in Introductory Geology Classes

Germán Mora[1]

ABSTRACT

Although active learning methodologies have been implemented in geoscience classes successfully, no direct comparison between these different instructional techniques exists to date. For that reason, the purpose of this study was to compare the effectiveness in student learning of two active learning methods: peer instruction and lecture tutorials. In particular, this study focuses on a first implementation of these active learning teaching methods in small- to medium -size introductory physical geology classes. Evaluation of their effectiveness was measured through the Geoscience Concept Inventory, which was administered at the beginning (pre-test) and at the end (post-test) of each course. In addition, students were asked to evaluate the contribution of these techniques to their own learning using a Likert like survey. A comparison of pre- and post-test results indicates that both methods provided statistically significant cognitive knowledge and understanding gains. A comparison of the post-test results for both methods reveals no statistical distinction, indicating a similar level of effectiveness for both peer instruction and lecture tutorials. Similarly, the vast majority of students indicated that these teaching techniques were instrumental in helping them learn different geologic concepts. The combined results of this study are consistent with others studies showing an improvement in cognitive knowledge and understanding gains whenever active learning instructional techniques are first implemented in science classes. A detailed analysis of the obtained data revealed that most of the gains were made by students having little prior knowledge of geology relative to those having some prior knowledge of geologic concepts. Given the relatively easy use of these techniques, their proven effectiveness, and the recognition by students of their effectiveness, it is then recommended that a wider implementation of these techniques should be used in introductory geology classes.

INTRODUCTION

College introductory science classes are typically the last opportunity that most students ever have to learn science in a formal setting. For that reason, these classes provide unusual challenges to instructors, given the variety of potential class outcomes, the number of topics, the range of expectations that students have for the class, their prior conceptual knowledge, their attitude towards science, and their diversity in learning styles (Bransford et al., 2000). Besides these challenges, surveys indicate that while instructors tend to overestimate pupils' learning gains (Rovick et al., 1999; Diakidoy and Iordanou, 2003), students often struggle with course vocabulary and are typically unprepared to develop successful strategies to learn the material (Lee et al., 1995; McCarthy and Kuh, 2006). These characteristic struggles, coupled with traditional pedagogical methods that normally lead instructors to misjudge students' understanding ultimately result in unchanged misconceptions about science and in little gain in conceptual knowledge. For example, it is well documented that traditional teaching methods produce small, negligible, or even negative gains in students' knowledge and understanding of scientific concepts (Halloun and Hestenes, 1985; McDermott, 1990; Hake, 1998; Libarkin and Anderson, 2005). Part of the problem is associated with the tendency of students in traditional teaching classes to resort to memorization as their strategy to learn the material in response to a learning environment that does not foster conceptual understanding (Tobias, 1990, 1992).

To increase students' conceptual and knowledge understanding and to improve their attitude towards science, a consensus exists about the need to develop novel teaching strategies that are based on a new teaching paradigm in which the class dynamic is fundamentally altered from a teacher-centered to a student-centered approach (Johnson et al., 1991; National Research Council, 1997). In this new class dynamic, students engage in their own learning process, which has produced both better attitudes towards science (Reynolds and Peacock, 1998) and more accurate conceptual understanding (Crouch and Mazur, 2001; Prather et al., 2005; Lasry et al., 2008). These student-learning methods involve students discussing questions and solving problems in class, with a significant amount of work being done by students working in groups. The success of these active learning strategies (*sensu* Bransford et al., 2000) appears to be related to the notion that they allow new information to be fitted into existing cognitive structures, which promotes learning. In contrast, when students perceive no apparent connections of new information to their prior knowledge and beliefs, as it is typically the case in teacher-centered classes, the new information is memorized and discarded (Bransford et al., 2000).

The dilemma for instructors is to determine the type of instructional method that would be best beneficial in her/his class, considering that a number of active learning strategies exist. Two strategies, in particular, have been shown to be useful in large introductory science classes: peer instruction (PI) and lecture tutorials (LTs). The main advantage of these instructional techniques from the instructor's point of view is that they are relatively easy to implement since they flow easily with traditional lectures, which are the most familiar type of instruction to college instructors. From the learners' perspective, these techniques have been shown to produce statistically measurable gains in concept understanding (Crouch and Mazur, 2001; Prather et al., 2005; McConnell et al., 2006; Lasry et al., 2008; Kortz et al., 2008). Both PI and LTs are approaches that improve learning because both involve active and cooperative learning, allow rapid feedback,

---

[1] Environmental Studies Program, Goucher College, 1021 Dulaney Valley Road, Baltimore, MD 21214; german.mora@goucher.edu

promote the development of students' higher-order thinking skills, and help students develop a disciplinary knowledge base. Although these two active learning strategies improve student learning, no study to date has compared their effectiveness directly. Thus, the purpose of this research study is to compare students' conceptual gain using these two approaches to establish whether a significant difference exists between these two teaching techniques.

PEER INSTRUCTION AND LECTURE TUTORIALS

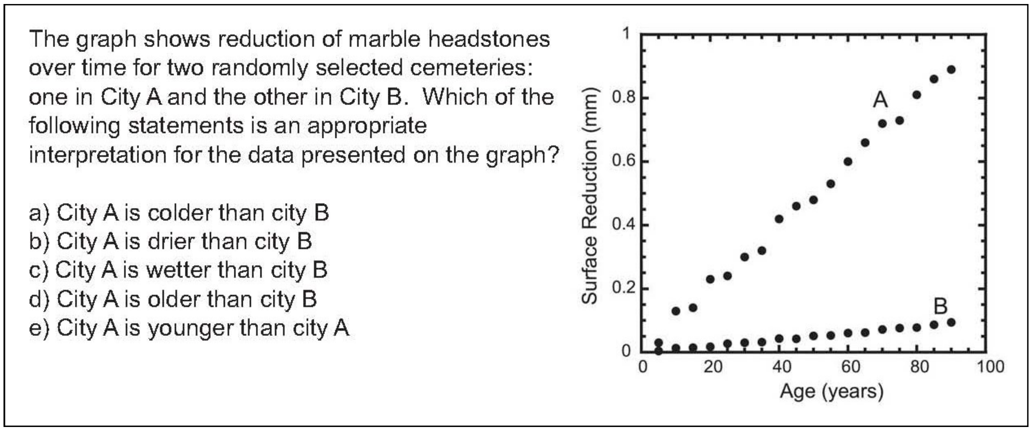
PI is an instructional method initially developed for physics classes over 10 years ago (Mazur, 1997). PI relies on a short explanation of a concept by the instructor, followed by a higher-order multi-choice question (or ConcepTest) that assesses students' understanding of the concept that was presented. Upon reporting the answer to the instructor, students are encouraged to explain their answers to other students and to re-think the multi-choice question. Students again report their answers to the instructor, who then moves to the next concept, if the answers provided by most of the students are correct. Otherwise, the instructor explains the concept using a different approach or provides additional examples until most students arrive at the correct answer. The explanation and negotiation that takes place between students typically results in a deeper understanding of the concept. In fact, several studies have demonstrated the effectiveness of this instructional method, since it typically results in higher students' concept understanding (e.g., Crouch and Mazur, 2001; McConnell et al., 2006; Lasry et al., 2008). PI has been expanded beyond college physics classes to chemistry (Wimpfheimer, 2002; Donovan, 2008), mathematics (Schlatter, 2002), chemical engineering (Falconer, 2007), and astronomy (Green, 2003). PI has also been employed in geology classes (McConnell et al., 2006). By studying nine different introductory geology classes and ~630 students, McConnell et al. (2006) found measurable gains in students' understanding of geologic concepts, improved attendance to class, and increased satisfaction. Moreover, instructors of these nine classes reported that implementation of PI in their courses was uncomplicated.

LTs are worksheets designed to use students' preconceptions and challenge typical misconceptions about a given topic by resolving conflicts with nature, logic, and consistency (Posner et al., 1982; McDermott, 1991, 1993). Typically, an instructor introduces a concept with a mini-lecture, poses a higher-order question to evaluate students' understanding of the topic, and provides the tutorials to students so they can solve it in small groups if their response to the question was inadequate. Once the tutorial is completed, the instructor either poses an additional question or explicitly provides an explanation to understand the concept. The apparent success of tutorials is related to the cognitive scaffolding and cognitive conflict that results when students work on the tutorial. The tutorials involve a series of progressive tasks and Socratic-type questions that are written in lay language and that include the analysis and interpretation of tables, figures, sketches, diagrams, or data. Students are asked to discuss and write their answers. To address common misconceptions, students are encouraged to evaluate a hypothetical "student debate" (McDermott et al., 1998), whereby two hypothetical students provide their explanations or conclusions of the material covered in the tutorial. While one hypothetical student provides a scientifically valid explanation, the other offers a common misconception. By working on the tutorials, students gradually improve their understanding of a given topic to the point that they are forced to confront their preconceptions, ultimately resulting in a more accurate understanding of a concept. For instance, several studies have documented that this instructional method promotes student learning (Prather et al., 2005; Mauka and Hingley, 2005; Kortz et al., 2008). The implementation of LTs in introductory geology classes has been successful as documented by Kortz et al. (2008), who evaluated their effectiveness in introductory geology classes consisting of about 86 students. These authors found that LTs produced measurable gains in students' understanding of geologic concepts, and the students enrolled in the classes found the LTs useful in learning the course material.

To evaluate whether the use of PI or LTs results in significant differences in learning and understanding of geological concepts by college students enrolled in an introductory physical geology class, conceptual understanding data from courses where the PI was implemented (Fall, 2008) were statistically compared to data from courses where LTs were implemented (Spring, 2009).

RESEARCH DESIGN AND METHODS

PI was implemented in the two daytime course sessions of Physical Geology (N=45) offered at Montgomery College, a public, two-year institution located in Rockville, Maryland, in the Fall of 2008. Implementation of this teaching approach followed the recommendations of Mazur (1997). Briefly, three to four ConcepTests were employed in each class session throughout the semester. Whereas some ConcepTest questions (~25%) came from the database developed by Steer and McConnell (2009), most of the questions (~75%) were developed by the course instructor following the guidelines of Mazur (1997) and McConnell et al. (2006). Figure 1 provides an example of a ConcepTest question that was employed to assess student's understanding of weathering rates. LTs were implemented in the two daytime course sessions of the same class in the Spring of 2009 (N=37). Implementation of this teaching approach followed the recommendations of Prather et al. (2005) and Kortz et al. (2008). Briefly, a tutorial was employed when needed, which was almost every other class meeting. The class met for 75 minutes twice a week for fifteen weeks. Whereas the majority (8) of the tutorials were developed by Kortz and Smay (2010), six tutorials were developed by the course instructor following the guidelines of Kortz et al. (2008). Appendix 1 provides an example of a LT that was employed to facilitate a better understanding of weathering rates. The same instructor taught both the PI and LTs sections, and the instructor had no prior


The graph shows reduction of marble headstones over time for two randomly selected cemeteries: one in City A and the other in City B. Which of the following statements is an appropriate interpretation for the data presented on the graph?

a) City A is colder than city B
b) City A is drier than city B
c) City A is wetter than city B
d) City A is older than city B
e) City A is younger than city A

FIGURE 1. Example of a ConcepTest employed in the study.

experience implementing either approach other than taking workshops on how to apply these two instructional methods. As a result, this study evaluates the effectiveness of a first implementation of these teaching techniques.

All course sessions covered the same concepts, employed the same textbook, and followed the same outlines, pace, tests, and lab assignments. No major differences in the distribution of age, years of college experience, percentage of non-science majors, and percentage of students who dropped were observed in the student population enrolled in the sessions (Table 1). The only difference was in gender, with the PI class sessions having more female students relative to the LTs class sessions (Table 1).

Student learning of basic geologic concepts was measured through the Geoscience Concept Inventory (GCI) (Libarkin and Anderson, 2005). Validation and reliability of the GCI have been determined through both qualitative and quantitative means, including classical test theory and Rash-based approaches (Libarkin and Anderson, 2005). The GCI v. 1.0 consists of a subset of validated questions that are distributed in eleven different difficulty levels. To allow comparisons with results from other studies, the GCI has four anchor questions, plus a question from each of the eleven difficulty levels. The GCI was given to students in both the PI and LTs classes during the second class meeting (pre-test) and the second-to-last class meeting (post-test) of each semester. To measure the effectiveness of the two teaching techniques, pre- and post-test scores were compared statistically, and any difference was considered significant when the p values were lower than 0.05. All statistical analyses were performed using Minitab 15. To evaluate the effectiveness of the two classes in terms of students' geologic concept understanding as measured by the GCI scores, the distribution of post-test scores was compared to that of pre-test scores to determine any significant difference between these two classes. In addition, students were asked to assess their experience in using both PI and LTs using a Likert type items on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree), which was given during the last class meeting. The survey included statements that prompted students to reflect on their own learning. Whereas validity of the statements of the survey was evaluated by three experienced instructors that have positive views about active learning, reliability was assessed only internally through the correlation that was obtained between paired statements placed in the survey that included positive and negative views of the same construct (see Table 2). Items were dichotomously scored

TABLE 1. STUDENT DEMOGRAPHICS

	Peer Instruction	Lecture Tutorials
Age	22.0±4.0[1]	20.7±3.5
Years of college[2]	1.8±0.7[1]	1.5±0.6[1]
Non-science majors[3]	91%	89%
Race	57% Caucasian	62% Caucasian
Gender	74% F, 26% M	47% F, 53% M
Students who dropped	8%	5%
Final class size[4]	45	37

[1] Standard deviation
[2] Students reported the number of semesters of full-time college enrollment. Part-time students estimated the fraction of full-time enrollment and provided the sum of these fractions.
[3] Undeclared majors were assumed to be non-science majors.
[4] Number of students whose pared pre- and post-test scores are available.

TABLE 2. EXAMPLE OF STATEMENTS INCLUDED IN THE LIKERT-TYPE SURVEY

Statements[1]
Peer instruction decreased my interest toward geology.
Peer instruction helped me better learn different geologic concepts.
Peer instruction decreased my overall understanding of geology.
Peer instruction made my learning experience more enjoyable.

[1] Peer instruction was replaced by Lecture tutorials in the survey provided to the other group.

(0 for positive views, 1 for negative views), and the consistency of responses was evaluated by the Kuder-Richardson coefficient. The obtained reliability coefficient for the statements was 0.92, which suggests internal consistency of the obtained answers.

To evaluate the validity and reliability of the entire study, the transferability, dependability, and confirmability (Lincoln and Guba, 1985) of its results are discussed here. Transferability addresses the applicability of the results and findings of the study in other contexts, such as other introductory geology courses. Table 1 provides information on the student population, indicating that it corresponds to traditional non-science students. This information and a detailed study design that includes the use of the GCI allow other researchers to apply the study to other contexts. To evaluate transferability, however, the GCI results obtained in this study from a relatively small population are compared to those obtained from a larger population (Libarkin and Anderson, 2005). Dependability addresses the repeatability of the study. To evaluate it, the GCI data from this study are compared to results from other studies evaluating the effectiveness of PI and LTs. Finally, confirmability addresses the objectivity of the study, which was initially evaluated through statistical comparisons of the produced data, followed by an external evaluation through the peer review process. Moreover, a detailed study design and the results are included here to permit a confirmation of the findings by other researchers.

RESULTS

Results from the GCI were scaled as recommended by Libarkin and Anderson (2006). Figure 2 shows box plots of the distribution of pre- and post-test scores for both PI and LTs classes, and Table 3 shows the mean and standard deviation of pre- and post-test (GCI) scores for both the PI and the LTs classes. The Kolmogorov-Smirnov test suggests that the data are marginally normally distributed. With this assumption of normality, paired t-test results are reported here whenever two sets of data were compared. However, non-parametric tests (e.g., Mann-Whitney Test) were also performed to compare the same two sets of data to identify discrepancies with the t-test results. In all cases, significant differences or similarities between two data sets deduced from the non-parametric tests were identical to those provided by the t-tests. The box plots illustrate an overall shift toward higher post-GCI scores for both classes relative to their pre -GCI scores (Fig. 2). The same figure illustrate that that there is no major difference between the pre-GCI scores for the PI and LTs classes. The same is true for their post-GCI scores. Statistical analyses of the dataset confirm the comparisons observed in Figure 2. Whereas the mean pre-test score for the PI class was 38.2 (SD=13.0), the mean for the LTs class was 36.4 (SD=12.1). There was no statistical difference between these two mean pre-test scores ($t = 0.70$, $p = 0.483$). The mean post-test scores for the PI and LTs classes were 49.7 (SD=12.3) and 46.9 (SD=13.8), respectively, and these mean values are statistically undistinguishable ($t = 1.02$, $p = 0.310$). Both post-test scores were higher than their respective pre-test scores ($t = 4.22$, $p < 0.0001$ for the PI class, and $t = 3.69$, $p < 0.001$ for the LTs class). A statistical comparison between the scores for the PI and LTs classes reveals no significance difference ($t = 0.26$, $p = 0.7$) between the two instructional methods.

Because LTs were not employed in every class and because some of the concepts evaluated by the GCI questions were indirectly covered by the LTs, a possible undue comparison may exist when contrasting GCI results for these two classes. To validate the findings from the GCI results, concepts that are tested in the GCI and that were directly discussed in both the LTs and PI classes were selected for comparison. Only two GCI questions met these criteria, namely global distribution of volcanoes and internal layering of the planet, and the percentages of

TABLE 3. PRE-GCI (GCI_{pre}) AND POST-GCI (GCI_{post}) SCORES AND STANDARDIZED GAINS (g)

	N	GCI_{pre}	GCI_{post}	GCI_{post} $-GCI_{pre}$	g
PI[1]	45	38.2 (SD=13.0)	49.4 (SD=12.3)	11.2	0.18
LTs[2]	37	36.4 (SD=12.1)	46.9 (SD=13.8)	10.5	0.16
Libarkin & Anderson[3]	2493[4]	41.5 (SD=12)	45.8 (SD=13)	4.3	0.07
Kortz et al.[5]	86[6]	39.3 (SD=13)	48.0 (SD=14)	8.7	0.14

[1] Peer Instruction classes reported in this study
[2] Lecture Tutorials classes reported in this study
[3] Mostly traditional classes reported (2005)
[4] The number of students with available post-GCI scores 1498
[5] Lecture tutorials classes reported (2008)
[6] The number of students with available post-CGI scores is 64

correct answers for these two questions between PI and LTs classes were compared. For the global distribution of volcanoes, students in the PI class improved from 20% correct in the pre-test to 35% correct in the post-test. Similarly, students in the LTs class improved from 22% to 36% correct. For the internal layering of the planet, PI students improved from 17% to 58% correct, whereas students in the LTs class improved from 14% to 54% correct. This comparison also indicates that a similar level of improvement was achieved in both classes.

Results from the surveys indicate that whereas 97% of the students enrolled in the PI class reported that they were satisfied with this methodology, 94% of the students enrolled in the LTs class were satisfied with this methodology. Similarly, 89% of the PI students and 87% of the LTs students indicated that they learned the concepts through these methods. Results from the survey also indicate that 83% of the PI students and 86% of the LTs students considered that these active learning techniques made their learning experience more enjoyable.

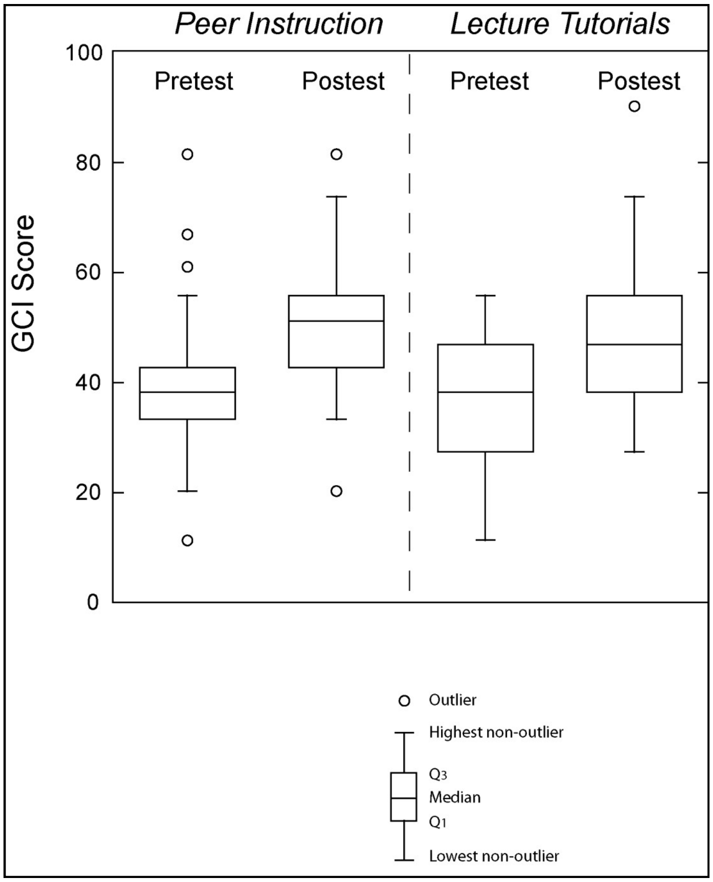


FIGURE 2. Box plots showing higher post-GCI scores for students in the Peer Instruction and Lecture Tutorials classes relative to their pre-GCI scores. Box plots provide a graphic representation of the distribution of data, with the lower and upper boundary of the box representing the first quartile (Q1) and the third quartile (Q3), respectively, the line within the box representing the median of the data set, and the two vertical lines from the box or "whiskers" extending to the lowest and highest non-outlier values in the data set. The circles represent outliers, which are values in the data set that are at least 1.5*IQR (Inter-Quartile Range = Q3 – Q1) beyond the quartiles. Notice that while overall post-GCI scores are higher relative to pre-GCI scores for both the PI and LTs classes, post-GCI scores for the PI class are comparable to those for the LTs class. The same is true for pre-GCI scores.

DISCUSSION

A comparison of pre- and post-GCI scores reveals that both PI and LTs promoted learning given that the post-test scores are significantly higher relative to the pre-test scores by 11.2 and 10.5, respectively. A comparison of these scores to those reported in other studies could provide some indication about the performance of the students enrolled in both classes relative to their peers. This comparison is effective only if the same type of GCI test was employed.

McConnell et al. (2006) reported the results of implementing PI in introductory geology classes, finding that whereas pre-GCI scores ranged from 41 to 48, post-GCI scores ranged from 49 to 58. These authors found that post-GCI scores were higher by 8 to 12, relative to their respective pre-GCI scores. Although not explicitly stated, it appears that these authors employed the GCI v. 1.0, which would allow a direct comparison with the results obtained in this study. The results reported by McConnell et al. (2006) are comparable to the 11.2 difference in pre- and post-GCI scores observed in this study. Moreover, a study of the use of LTs in introductory geology classes (Kortz et al., 2008) found an increase in GCI scores from 39.3 (SD=13, n=86) to 48.0 (SD=14, n=64). Although not explicitly stated, these authors probably employed the GCI v. 1.0, which would allow a direct comparison with the results obtained in this study. The average increase of 8.3 reported by Kortz et al. (2008) is slightly lower than the 10.5 found in this study, but the comparisons between the results of this study and those of other studies that employed either PI or LTs in introductory geology classes corroborate the effectiveness of these two teaching techniques. Their effectiveness is also illustrated by a comparison of the gains observed in the PI and LTs classes relative to the gains reported for ~2,500 students enrolled in mostly traditional introductory geoscience courses across the United States (Libarkin and Anderson, 2005). Results from this and other studies (McConnell et al. 2006; Kortz et al., 2008) indicate that students in either PI and LTs classes obtained consistently higher GCI scores than those in the 30 classes surveyed by Libarkin and Anderson (2005), who found a statistically significant improvement in post-GCI scores relative to pre-GCI scores in only 13 out of the 30 introductory geoscience courses surveyed.

One way to evaluate the effectiveness of instructional methods as measured by tests like the GCI is to express the results in terms of average normalized gains (Post-test scores - Pre-test scores/100 - Pre-test scores), which provides a readily accessible, objective measure of learning (Hake, 1998). The results from this study reveal that the average normalized gains almost double relative to those reported for ~2500 students enrolled in the 30 classes surveyed by Libarkin and Anderson (2005) (Table 3), which is consistent with results indicating that there is a significant improvement in cognitive gains whenever active learning instruction is involved (Hake, 1998; Crouch and Mazur, 2001). Notably, Table 3 also shows that the standardized gains for the LTs classes in this study and those for the LTs study by Kortz et al. (2008) are nearly identical. Unfortunately, standardized gains for the PI study by McConnell et al. (2006) cannot be calculated from the published data. The equal success of these two instructional methods could be associated with the underlying premises and similarities associated with both methods. Both methods rely on a mini-lecture of a concept and a higher-order question that assesses students' understanding of the concept. Learning occurs as both methods challenge students' misconceptions and employ a social context that promotes collaboration (Heller and Hollabaugh, 1992; Heller et al., 1992; Kalman et al., 1999).

An important question is whether all students benefited equally from these active-learning techniques. Figure 3 shows standardized gains relative to pre-GCI scores, revealing a negative correlation. This negative correlation suggests that not all students benefited equally from these active-learning techniques. To investigative this suggestion further, the student population was divided into those who performed below the median pre-test score (38.25) and those who obtained the median score or higher. Results from this division (Fig. 4) reveal that almost all the normalized gains were associated with students whose pre-test scores were below the median. In contrast, no gains were associated with students whose pre-test scores were equal or higher than the median score (Fig. 4). Students with below median pre-test scores in the PI class obtained statistically significant standardized gains in geologic concept understanding relative to students with median pre-test scores or higher ($t=2.91$, $p<0.005$). A similar finding was observed in the LTs class, but with less significance given the variance observed for students with above median pre-test scores. These results indicate that both PI and LTs chiefly benefited students with lower prior understanding of geologic concepts, but these instructional methods had no significant benefit on students with some prior understanding of geologic concepts. This finding, however, has not been reported for other classes that include the implementation of active learning instruction, which tends to produce significant gains in both groups (Lasry et al., 2008). At the moment, the available data from this study do not provide any conclusive evidence to explain this finding, but either the implementation of these techniques or the instrument to measure students' gains could potentially account for the observed data. One possible explanation is that the implementation of both PI and LTs was not fully achieved. Although all measures were taken to ensure that the implementation of these techniques followed the parameters and recommendations outlined in the literature, no expert evaluated the first implementation of these techniques by the instructor. As a result, no independent evaluation exists about the possibility that some aspects of these techniques might have not been entirely applied to the studied student population. Ruling out this possibility, however, the obtained results from this study are similar to those reported whenever active learning techniques are implemented in science classes (Hake, 1998; Crouch and Mazur, 2001; Kortz et al., 2008). Another possibility may involve the difficulty associated with some GCI questions. It is possible that some questions are significantly more difficult than others. As a result, it would be easier to reach pre-GCI median scores than it is to surpass those scores. Students with some prior knowledge of geologic concepts would have to invest significantly more effort to improve their pre-GCI scores, relative to those with little prior knowledge of geologic concepts. This possible "ceiling effect" embedded in the GCI would have to be verified with other studies.

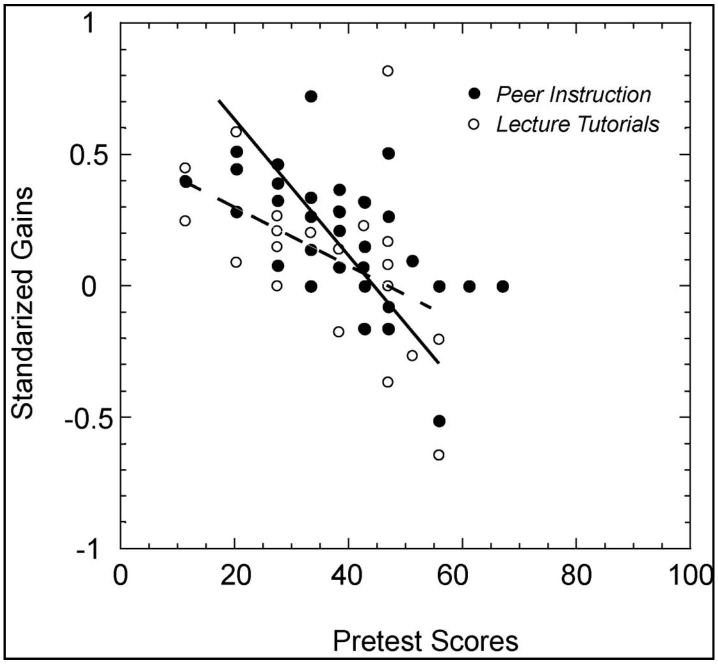


FIGURE 3. Cross plot showing a slight negative correlation between the distribution of standardized gains (Post-GCI scores - Pre-GCI scores/100 - Pre-GCI scores) and pre-GCI scores. Positive (negative) standardized gains indicate higher (lower) post-GCI scores relative to their respective pre-GCI scores. Whereas the solid line is the linear regression for the PI data ($y = 1.19 - 0.028x$, $r = 0.64$), the dashed line is that for the LTs data ($y = 0.55 - 0.011x$, $r = 0.48$).

There are a couple of caveats about the findings reported in this study that are important to consider. The sample size is relatively small, but the obtained results are consistent to those obtained in other studies with slightly larger sample size (e.g., McConnell et al., 2006; Kortz et al., 2008). Another caveat involves the possibility that the gains reported in this study may not entirely be due to the implementation of PI or LTs, since a number of factors are involved in learning. For example, the LTs covered less than half of the total number of concepts in the class. Nonetheless, the variables were minimized in this study to produce a fair comparison between the methods, and results of this study are consistent with the results of other studies dealing with the implementation of active-learning techniques in introductory science classes in general (Crouch and Mazur, 2001; Prather et al., 2005; Lasry et al., 2008) and in geoscience classes in particular (McConnell et al., 2006; Kortz et al., 2008). Moreover, a direct comparison employing only GCI questions of concepts that were directly covered by both instructional methods reveal similar level of improvement of correct answers.

The combined body of evidence strongly suggests that these active-learning teaching techniques are significantly beneficial to students enrolled in introductory science classes. Besides improving student learning, these techniques are also relatively easy to implement as has been explained elsewhere for both PI (McConnell et al., 2006) and LTs (Kortz et al., 2008). Their advantage from the instructor's perspective, is that instructors do not have to implement drastic changes in course content or organization, as both techniques require mini-lectures. Moreover, the time invested in implementing these techniques is not substantial since an existing database of ConcepTests and Lecture Tutorials is already available (Steer and McConnell, 2009; Kortz and Smay, 2010). Similarly, over three quarters of students expressed in surveys that these techniques helped them learn geologic concepts more effectively and made the course more enjoyable and less threatening.

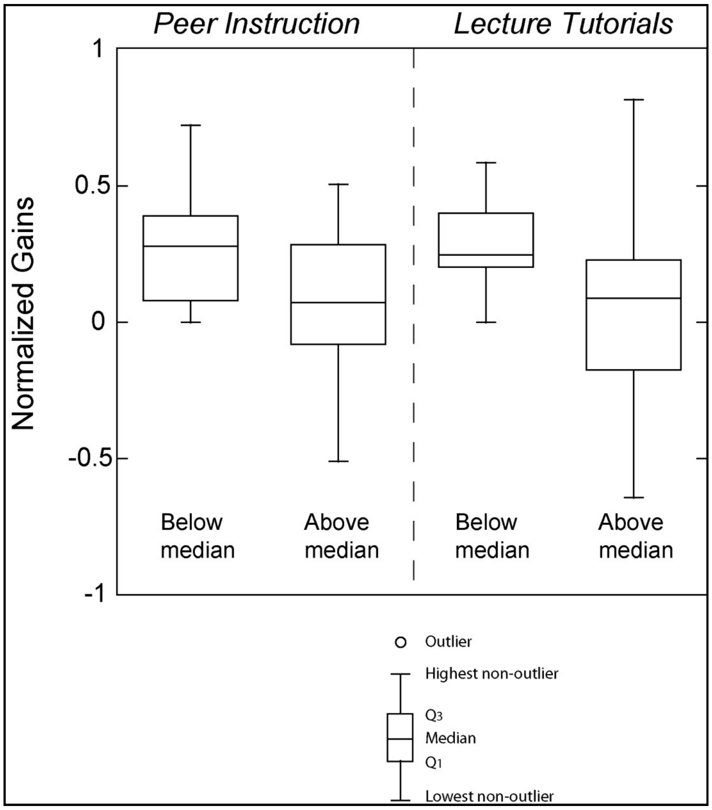


FIGURE 4. Box plots showing the distribution of standardized gains (Post-GCI scores – Pre-GCI scores/100 – Pre-GCI scores) for students whose pre-GCI scores are below and equal or above the median pre-GCI score (38.25) in the Peer Instruction and Lecture Tutorials classes. Positive (negative) standardized gains indicate higher (lower) post-GCI scores relative to their respective pre-GCI scores. The lower and upper boundary of the box represent the first quartile (Q1) and the third quartile (Q3), respectively, the line within the box represents the median, and the two vertical lines from the box or "whiskers" extend to the lowest and highest non-outlier values in the data set. Notice that students with pre-GCI scores below the median in both the PI and LTs classes have overall higher standardized gains than those with high pre-GCI scores.

CONCLUSION

Results from this study indicate that both PI and LTs increase student learning of geologic concepts by nearly doubling standardized gains relative to those reported in a large study of students enrolled across the United States (Libarkin and Anderson, 2005). The gains reported in this study are similar to those reported in other studies related to the first implementation of these active-learning techniques. Students report that these techniques helped them in their learning process and made the class environment more enjoyable. However, a more detailed analysis of the data reveals that most of the gains were associated with students with relatively lower prior knowledge of geologic concepts. In contrast, no gains in conceptual understanding were observed, on average, for students with relatively higher prior knowledge of geologic concepts. Despite this finding, these students overwhelmingly indicated that these instructional methods were instrumental in learning geologic concepts and enjoying the class.

Acknowledgments

The author would like to thank Dr. Julie C. Libarkin for her advice and comments on administering the Geoscience Concept Inventory, Dr. Alessandro Zanazzi for his suggestions on the statistical analysis of the data, the administration at Montgomery College for allowing me to carry out this study, and three anonymous reviewers and the editorial staff at JGE for providing thoughtful comments that improved the manuscript.

Appendix 1: Example of a lecture tutorial employed in the study (see three pages following References).

REFERENCES

Bransford, J.D., Brown, A.L., and Cocking, R.R., 2000, How People Learn: Brain, Mind, Experience, and School. Washington, D.C., National Academy Press. Online at http://www.nap.edu/books/0309070368/html/.

Crouch, C.H., and Mazur, E., 2001, Peer Instruction: Ten years of experience and results. American Journal of Physics, v. 69, p. 970-977.

Diakidoy, I.A.N., and Iordanou, K., 2003, Preservice teachers' and teachers' conceptions of energy and their ability to predict pupils' level of understanding. European Journal of Psychology of Education, v. 18, p. 357-368.

Donovan, W., 2008, An electronic response system and Conceptests in general chemistry courses. Journal of Computers in Mathematics and Science Teaching, v. 27, p. 369-389.

Falconer, J.L., 2007, Conceptests for a chemical engineering thermodynamics course. Chemical Engineering Education, v. 41, p. 107-114.

Green, P.J., 2003, Peer Instruction in Astronomy. Addison-Wesley, 178 p.

Hake, R. R., 1998, Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, v. 66, p. 64-74.

Halloun, I., and Hestenes, D., 1985, The initial knowledge state of college physics students. American Journal of Physics, v. 53, p. 1043-1055.

Heller, P., and Hollabaugh, M., 1992, Teaching problem solving through cooperative grouping. Part 2: designing problems and structuring groups. American Journal of Physics, v. 60, p. 627-636.

Heller, P., Keith, R., and Anderson, S., 1992, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. American Journal of Physics, v. 60, p. 627-636.

Johnson, D. W., Johnson, R. T., and Smith, K. A., 1991, Active Learning: Cooperation in the College Classroom, Edina, MN., Interaction Book Company. 140 p.

Kalman, C.S., Morris, S., Cottin, C., and Gordon, R., 1999, Promoting conceptual change using collaborative groups in quantitative gateway courses. Physics Education Research, American Journal of Physics Supplement, v. 67, p. S45–S51.

Kortz, K.M., and Smay, J.J., 2010, Lecture Tutorials in Introductory Geoscience. H.W. Freeman, 100p.

Kortz, K.M., Smay, J.J., and Murray, D.P., 2008, Increasing student learning in introductory geoscience courses using Lecture Tutorials. Journal of Geoscience Education, v. 56, p. 280-290.

Lasry, N., Mazur E., and Watkins, J., 2008, Peer instruction: From Harvard to the two-year college. American Journal of Physics, v. 76, p. 1066-1069.

Lee, O., Fradd, S.H., and Sutman F.X., 1995, Science knowledge and cognitive strategy use among culturally and linguistically diverse students. Journal of Research in Science Teaching, v. 32, p. 797-816.

Libarkin, J.C., and Anderson, S.W., 2005, Assessment of learning in entry-level geosciences courses: Results from the Geoscience Concept Inventory. Journal of Geoscience Education, v. 53, p. 394-401.

Libarkin, J.C., and Anderson, S.W., 2006, The Geoscience Concept Inventory: Application of Rasch Analysis to Concept Inventory Development in Higher Education: in Applications of Rasch Measurement in Science Education, ed. X. Liu and W. Boone: JAM Publishers, p. 45-73.

Lincoln, Y.S., and Guba, E.G., 1985, Naturalistic Inquiry, Beverly Hills, Sage Publishing, 416 p.

Mauka, H.V., and Hingley, D., 2005, Student understanding of induced current: Using tutorials in introductory physics to teach electricity and magnetism. American Journal of Physics, v. 73, p. 1164-1171.

Mazur, E., 1997, Peer Instruction: A User's Manual. Upper Saddle River, NJ, Prentice Hall, 253 p.

McCarthy, M., and Kuh, G. D., 2006, Are students ready for college? What student engagement data say. Phi Delta Kappan, v. 87, p. 664-669.

McConnell, D.A., Steer, D.N., Owens, K., Borowski, W., Dick, J., Foos, A., Knott, J.R., Malone, M., McGrew, H., Van Horn, S., Greer, L., Heaney, P.J., 2006, Using ConcepTests to assess and improve student conceptual understanding in introductory geoscience courses. Journal of Geoscience Education, v. 54, p. 61-68.

McDermott, L. C., 1990, A view from physics. In M. Gardner et al. (Ed.), Toward a scientific practice of science education (pp. 3–30). Hillsdale, NJ: Erlbaum.

McDermott, L. C., 1991, What we teach and what is learned – Closing the gap. American Journal of Physics, v. 59, p. 301-315.

McDermott L. C., 1993, Guest comment: How we teach and how students learn – a mismatch? American Journal of Physics, v. 61, p. 295-298.

McDermott, L.C., Shaffer, P.S., and the Physics Education Group at the University of Washington, 1998, Tutorials in Introductory Physics. Upper Saddle River, NJ, Prentice Hall.

National Research Council, 1997, Science Teaching Reconsidered, Washington, D.C., National Academy Press, 88p.

Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A., 1982, Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, v. 66, p. 211.

Prather, E.E., Slater, T.F., Adams, J.P., Bailey, J.M., Jones, L.V., and Dostal, J.A., 2005, Research on a lecture-tutorial approach to teaching introductory astronomy for non science majors. Astronomy Education Review, v. 3, p. 122-136.

Reynolds, S.J., and Peacock, S.M., 1998, Slide observations – Promoting active learning, landscape appreciation, and critical thinking in introductory geology courses. Journal of Geoscience Education, v. 46, p. 421-426.

Rovick, A.A., Michael, J.A., Modell, H.I., Bruce, D.S., Horwitz, B., Adamson, T., Richardson, D.R., Silverthorn, D.U., and Whitescarver, S.A., 1999, How accurate are our assumptions about our students' background knowledge?. Advances in Physiology Education, v. 21, p. 93-101.

Steer, D., and McConnell, D., 2009, ConcepTests for introductory geoscience courses. Online at: http://serc.carleton.edu/introgeo/interactive/ctestexm.html.

Schlatter, M.D., 2002, Writing Conceptests for a Multivariable Calculus class. Primus, v. 12, p. 305-314.

Tobias, S., 1990, They're not dumb, they're different: Stalking the second tier, Tucson, Research Corporation, 94 p.

Tobias, S., 1992, Revitalizing undergraduate education: Why some things work and most don't, Tucson, Research Corporation, 192 p.

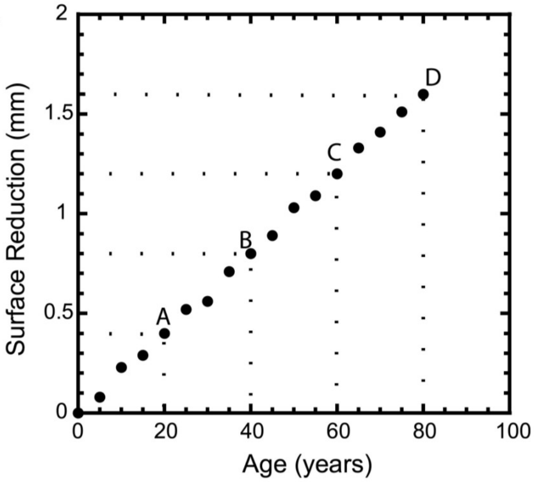
Wimpfheimer, T., 2002, Chemistry Conceptests: Considerations for small class size. Journal of Chemical Education, v. 79, p. 592.

APPENDIX 1. Example of a lecture tutorial employed in the study (page 1)

Weathering Rates

1. The graph shows data of how much the surface of marble headstones has been worn away since the time it was put in place in graveyards of a European city.

   What general trend do you observe?



2. One could place a line that passes through the middle of the points on the graph to represent the trend and roughly calculate how much the headstone wears away per year, which is called weathering rate. For this calculation, two random points are selected along the line, and their respective ages and surface reduction are estimated from the graph. See the illustration below for an example.

	B	**A**	**B – A**
Surface Reduction	0.8 mm	0.4 mm	0.4 mm
Age	40 years	20 years	20 years

Weathering Rate = Surface Reduction (B – A) / Age (B – A) 0.4/20 = 0.02 mm/year

Following this example, calculate the weathering rate for points C and D.

	D	**C**	**D – C**
Surface Reduction			
Age			

Weathering Rate = Surface Reduction (D – C) / Age (D – C)

3. Relative to points A and B, is the weathering rate for points C and D. (circle one)

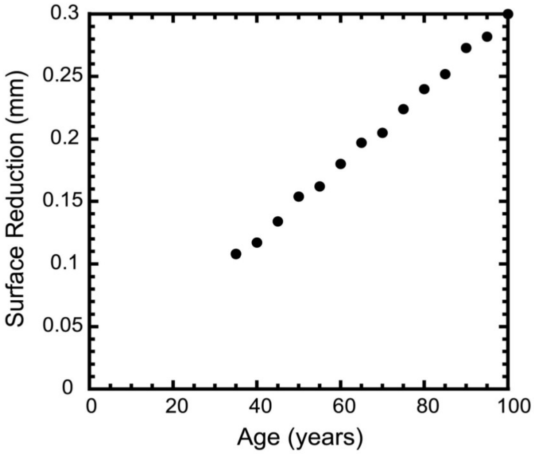
   The same   Higher   Lower

Weathering Rates

4. The graph shows data for headstones made of well-cemented sandstones from the same graveyards of the English city. Notice that the scale of the vertical axis in this graph is different relative to the graph above.

   Relative to the marble headstones in the example above, what is the weathering rate for these sandstone headstones? (circle one)

   The same Higher Lower



5. Briefly explain your answer.

6. Two students are debating about the weathering rates of the material of the headstones.

   Student 1: I see that the sandstone headstones tend to be older than the marble headstones, so I think that the sandstone headstones weather more quickly.
   Student 2: I don't think that the age of the headstones is related to weathering rate. For any time interval, I see that the sandstone headstones have weathered less than the marble headstones, meaning that the sandstone headstones weather more slowly. So sandstones have a lower weathering rate.

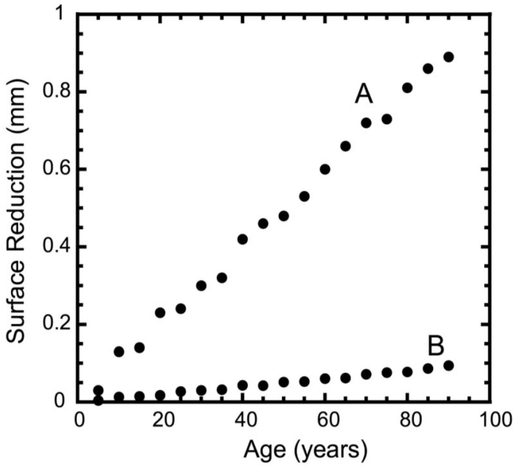
   With which student do you agree? Explain your answer.

7. Examine the figure showing data for marble headstones from two different cities. City A receives abundant rain, whereas City B receives little rain.

   In which city do you infer a higher weathering rate? (circle one)

   City A City B

   Explain your answer.



APPENDIX 1. Example of a lecture tutorial employed in the study (page 3)

Weathering Rates

8. Using the graph, determine the surface reduction at 30 and 60 years for both cities. With the data, then calculate the weathering rate for City A and for City B.

	60 Years	30 Years
Surface Reduction for City A		
Surface Reduction for City B		

Weathering Rate for City A		
Weathering Rate for City B		

9. Compare the weathering rates of City A and City B. Do they match your answer to Question 8?

10. In what climate would you expect a headstone to weather faster? (Circle one)

A dry climate A rainy climate

11. Let's say that you are a billionaire wanting to build a grand mausoleum for you and your family so it lasts for hundreds of thousands of years (preferably millions of years). What construction material and place in the world would you select to fulfill this dream? Explain your answer.