The Importance of a Laboratory Section on Student Learning Outcomes in a University Introductory Earth Science Course

Frank L. Forcino[1,a]

ABSTRACT

Laboratory sections of university Earth science courses provide hands-on, inquiry-based activities for students in support of lecture and discussion. Here, I compare student conceptual knowledge outcomes of laboratory sections by administering an independent concept inventory at the beginning and end of two courses: one that had a lecture and a laboratory section, and one that had only a lecture. Students in both courses demonstrated a significant increase in inventory scores over the course of the semester. The mean increase in score for the course with a laboratory was 33% greater (43% greater for matched-identification score analysis) than for the lecture-only course. One notable difference between the two courses was that the course without the lab was also a time-shortened course, while the course with a lab spanned a full, traditional-length academic term. Because a great deal of research exists demonstrating that time-shortened, intensive university courses produce the same increase in student concept knowledge as traditional-length courses, the inclusion of the laboratory section most likely led to the greater student learning gains in the full-length course. This study demonstrates the importance of having a laboratory component of an introductory-level, university Earth science course. © *2013 National Association of Geoscience Teachers*. [DOI: 10.5408/12-412.1]

Key words: time-shortened, intensive, Geoscience Concept Inventory

INTRODUCTION

Laboratory activities and experiments often accompany lecture and discussion portions of science courses. University science courses often have a separate laboratory section for students, in addition to a lecture section. The importance of the laboratory sections is argued by almost all science educators, using mostly anecdotal evidence to support their claims (Blosser, 1983; Baird, 1990; Hofstein and Lunetta, 2004). However, there has been little direct research into the effect of laboratories on student learning outcomes in university science courses (Doran, 1978; Hofstein and Lunetta, 1982, 2004; Prades and Espinar, 2010; Matz et al., 2012).

Specifically for the Earth sciences, laboratory sections of university courses are important for student learning (Neilson et al., 2010; Huysken et al., 2011). The laboratories ostensibly provide students with hands-on activities in which inquiry, experimentation, and discovery play key roles. In addition, laboratories are where student are able to spend time interacting with samples (e.g., rocks, minerals, and fossils) and conduct activities using maps and geographic information systems. Earth sciences are historically field-based sciences; geoscientists must be able to identify and interpret material in the field. Laboratories prepare students for future field-based experiences, as well as for more advanced Earth science courses. In addition, students learn material better when multiple modalities are employed (Penney, 1989; Cowan, 1998; Gadt-Johnson and Price, 2000; Prain and Waldrip, 2006). Because a laboratory section often adds tactile and visual learning methods for students that complement the auditory component from the lecture, it is important to determine whether the additional modalities lead to higher learning gains.

Independent concept inventories (validity- and reliability-tested assessments administered by someone not directly instructing students and not counted toward student grades) of university courses are becoming more prevalent (Libarkin and Kurdzel, 2001; Perkins, 2004; Libarkin and Anderson, 2006; Jolley et al., 2012). Because such inventories have low stakes, students may not put forth an optimal effort. However, in conjunction with course exams and projects, concept inventories inform instructors as to whether students understand the material and are accomplishing the desired learning outcomes. Graded activities can add confounds, such as curving of exams scores, which make it difficult to compare student performance from semester to semester. Variation in student attitude or aptitude from term to term also makes it problematic to use exam or course scores for comparing student performance in subsequent semesters. Independent concept inventories can remove possible biases (e.g., favoritism for a student) that the course instructor might unknowingly add, because the independent inventory is blinded.

Here, I compare student conceptual knowledge outcomes of laboratory sections by administering a modified version of the Geoscience Concept Inventory (GCI) at the beginning and end of two courses: one that had a lecture and a laboratory section, and one that only had a lecture. Two aspects that set this study apart from previous work are (1) a consistent independent concept inventory was used at the beginning and end of each course and (2) the same instructors taught the same material in both courses. Because one course had a laboratory section and one did not, any difference in the change in inventory scores is likely attributable to the laboratory, or lack thereof. My hypothesis was that the course with the laboratory section would lead to a greater increase in concept inventory scores from the

Received 22 December 2012; revised 28 January 2013; accepted 31 January 2013; published online 14 May 2013.

[1]Department of Earth and Atmospheric Sciences, University of Alberta, 1-26 Earth Sciences Building, Edmonton, Alberta, Canada, T6G 2E3
[a]Author to whom correspondence should be addressed. Electronic mail: flforcino@gmail.com. Tel.: 484-467-7887. Fax: 780-492-2030

TABLE I: Course attributes of both courses examined in this study.[1]

	Planet Earth (With Lab)	Earth Sciences I (Lecture Only)
Course objectives	Show how the Earth functions as a dynamic system that is shaped by interactions among its geological, physical, chemical, and biological elements. Overall, provide a solid foundation in all branches of Earth science.	Students will gain a basic understanding of the geoscience topics (see text), as they relate to the scientific method, global warming, global climate change, and interactions among the geosphere, atmosphere, hydrosphere, and biosphere, including the recent impacts of human activities.
Delivery	Lecture and lab	Lecture
Prerequisites	None	Any 100-level science course
Course length	13 weeks	3 weeks
Period length	50 min	2.5 h
Credit hours	3	3
Evaluation mode	Exams and lab exercises	Exams
GE requirements satisfied	Varies among programs	Varies among programs

[1]GE = general education.

beginning to the end of the course. This would corroborate previous research and anecdotal evidence that laboratory sections benefit student learning outcomes. To examine possible confounding effects because the lecture-only course was a time-shortened intensive course, I tested a subhypothesis that students enrolled in the course with the laboratory would score higher on questions that corresponded to content that was explicitly covered in at least one laboratory exercise. This would indicate that the laboratory exercises, rather than the difference between course lengths, were affecting student learning outcomes.

METHODS

Course Descriptions

I conducted this experiment in an Earth science department at a Canadian university. The department offers two versions of an introductory Earth science course (Table I). Planet Earth contains a laboratory section, while Earth Sciences I does not. Both courses offer an introduction to the origin and evolution of the Earth and solar system. Specific topics covered include plate tectonics, the rock cycle, energy balances, the water cycle, evolution of life, global climate change, the carbon cycle, human interaction with the Earth, and mineral and energy resources. The same lecture material was taught in both courses in each instructor's respective sections. The only apparent differences were (1) the lack of a laboratory section in Earth Sciences I and (2) Planet Earth was 13 weeks (with three 50-min lectures per week), whereas Earth Sciences I was 3 weeks (with 12.5 h of lecture per week). I was unable to test directly whether the lab or the full-length course caused differences in student learning gains in the lab course. However, previous research has found that intensive and traditional courses to be equally effective in producing the same student learning gains (Daniel, 2000, and references within; Davies, 2006, and references within; Kucsera and Zimmaro, 2010; Nasiri and Shokrpour, 2012). So, differences would likely be attributed to having the laboratory section or not.

The Planet Earth laboratory exercises support topics taught in the lecture section of the course. Laboratory activities include mapping and rock and mineral identification, and they cover topics such as the water cycle, glaciers, the atmosphere, the history of the Earth, and the human footprint on the Earth. Because the background information is taught in the lecture sections, and the lectures are the same for both the course with and the course without the lab, the students enrolled in the lecture-only course do not receive the hands-on experience of the laboratory to reinforce the lecture material (Table I).

Students may enroll in the lecture-only course if the general science requirements for their particular major do not include a science course with a laboratory (Table I). Often these students still enroll in the lab course because it is more conveniently scheduled during the semester and more sections are offered per year. Thus, there is a consistent mix of students from various majors among the courses. There is no evidence that a student more adept at science would enroll in one course instead of the other. However, I did test to determine whether a bias existed toward science majors in one course or the other (as explained later).

Instrument and Administration

The instrument used to evaluate the students was a modified version of the GCI. The GCI is a multiple-choice assessment instrument used in introductory geoscience courses. It consists of a set of 69 questions, from which I selected 28 (see the supplemental material, available at http://dx.doi.org/10.5408/12-412s1, for the complete concept inventory). Each of these questions has gone through rigorous reliability and validation studies (Libarkin and Anderson, 2005, 2006). These particular 28 questions were chosen from the entire pool of GCI questions because they most closely aligned with the lecture material. I constructed another nine questions to cover content that was not part of the GCI. These nine questions were created to measure content retention of material taught in the two courses. Incorrect response options were provided to reflect common misconceptions, based on statements from the instructors of the Planet Earth course. These nine questions are marked in red in the supplemental material.

The total score possible on the concept inventory instrument was 46 points—some of the 37 questions had multiple possible answers and, therefore, multiple possible

TABLE II: Dates each pre- and postcourse assessment was administered.[1]

	With Lab				Lecture Only	
Instructor	Precourse	Postcourse	Precourse	Postcourse	Precourse	Postcourse
Instructor 1	Sept. 9	Dec. 5	Jan. 11	Apr. 11	May 7	May 24
Instructor 2	Sept. 9	Dec. 5	N/A	N/A	May 28	June 13

[1]N/A = not applicable.

points. In addition to the 37 questions, students were asked to volunteer their student identification numbers in order to match scores from precourse to those from postcourse. The 37 questions were intended to gauge general Earth science knowledge. One advantage of using a standard instrument such as the GCI is that it is not specific to any university or to any particular course. The transferability of this instrument made it appropriate for this study.

I administered pre- and postcourse concept inventories to students between September 2011 and June 2012 (Table II). The concept inventory instrument was identical for both Earth and atmospheric science courses and for both the pre- and the postcourse inventories. The purpose of the precourse inventory was to measure students' incoming concept knowledge, while the postcourse inventory measured conceptual knowledge gained throughout the span of each course.

Student random guessing could lead to anomalously high scores. To determine the average assessment score that would be achieved by random guessing, I created a null model of randomly generated responses to the assessment questions. The null model was constructed by randomly selecting one response for each of the 37 questions. Although some questions had multiple correct answers (meaning that the test had 46 correct selections, for 46 total points), this aspect was omitted from the model. This simplified the model and took into account the likely scenario that students who ignored the instruction to not randomly select answers would also not pay attention to the test instructions, which directed students to select multiple responses when appropriate. The random selection of responses to all 37 questions was repeated 1,000 times to determine the mean score achieved by random guessing. This null model simulates 1,000 students randomly choosing responses to all assessment questions.

I obtained human subject research permission in accordance with applicable Canadian laws and university protocol. Student responses to inventory questions were coded to remove identifiers. Therefore, student anonymity was maintained even though identification numbers were collected, because I never obtained a list of student's names and I was the only person in possession of the scores. The variables I gathered from the students were instructor identity (one of two possibilities), class standing (first through fourth year), and academic major. For purposes of this study, I categorized the majors as Earth science, science, or nonscience. Earth science majors included geology, petroleum geology, Earth science, and atmospheric science majors. Science majors included other "hard" sciences, including chemistry, biology, ecology, and physics. Every other major was included as nonscience majors, including engineering and math majors. I did not collect student gender, age, or ethnicity because I was limited by class time; I was only allotted by each of the instructors a certain amount of time for students to take the concept inventory.

Statistical Analysis

To test for a statistically significant increase in student conceptual knowledge, I performed t-tests between pre- and postcourse inventory scores for the lecture-only course and the lab course, separately. Two sets of t-tests were conducted: (1) all scores from pre- and postcourse inventories for both the lecture-only course and the lab course were examined and (2) the individual scores—matched by student identification numbers—were sequentially compared to one another. For the first analysis, I conducted t-tests to test for significant differences between the pre- and the postcourse scores. For the second analysis, I conducted paired-sample t-tests to test for a significant difference between the two sets of inventory scores. The paired-sample t-tests take into account the changes of individual student's scores from pre- to postcourse.

In addition to the analysis that included all 37 questions, a secondary analysis was conducting examining 12 particular questions. These 12 questions contained content that was explicitly covered in at least one laboratory exercise. Because the concept inventory was initially designed for general assessment of the lab course, without this specific study in mind, the questions were geared toward general Earth science knowledge (the bulk of the material instructed in the lecture section). As a secondary test of the differences in student conceptual knowledge gains between the lecture-only and the lab courses, I compared the results of these 12 inventory questions (totaling 16 points) between the two courses. Because students in the lab course had a greater opportunity to actively engage in the material covered by these 12 questions, it is possible that the students in the lecture-only course would not demonstrate as strong of an increase in scores for these 12 questions. Therefore, evidence for a greater increase in the conceptual knowledge for these 12 questions among students enrolled in the lab course would provide evidence that the laboratory course directly aids in student learning outcomes.

Because 9 of the 37 questions (10 out of the total 46 points) on the inventory were not from the GCI and were not tested for reliability or validity, additional t-tests were conducted comparing standardized scores for only the 28 GCI questions to scores of all 37 questions. If there are no significant differences between the scores of the 28 GCI questions and all 37 inventory questions, this suggests that the 9 questions were not biased.

Within each course, there were two possible instructors, and the enrolled students pursued a range of majors, including Earth science majors, science majors, and nonscience majors. Furthermore, as is often the case with time-shortened, intensive courses, students in the lecture-only course tended to be further along in their university careers

TABLE III: The percentage (and number) of each year of student academic standing for both the course with the laboratory section and the lecture-only course.

	With Lab	Lecture Only
1st year	57% (122)	2% (1)
2nd year	26% (55)	14% (9)
3rd year	12% (25)	33% (21)
4th year	6% (12)	51% (32)

than those in the full-length course (Daniel, 2000; Table III). To evaluate the possible effects of instructor, academic major, and class standing on the change in inventory scores from the pre- to postcourse, I performed and analyzed general linear models (GLMs). The GLMs statistically evaluate the goodness of fit of the different populations (i.e., pre- and postcourse inventories for lecture-only and lab courses), different instructors, year, and major to the inventory scores. The GLM consisting of just the inventory scores for the pre- and postcourse populations is the same as a *t*-test. However, by conducting this univariate GLM, a baseline is created to compare to the bivariate GLM. To measure the effect of the instructor, year, and major on the change in concept inventory scores, I calculated the difference between the r^2 statistics between the univariate GLM (the two populations) and the bivariate GLM (the two populations and the different instructors, year, or major). In addition to the three bivariate GLMs, I conducted one multivariate GLM consisting of the inventory scores, year, and major. If any variable were to have a significant effect on the change in scores, the r^2 statistic would be greater in the bivariate analysis than the univariate analysis. However, in most cases, adding variables to the GLM would result in some increase in the r^2 statistic. So, each GLM was evaluated using the Akaike information criterion (AIC) and Bayesian information criterion (BIC; also known as the Schwarz criterion; Akaike, 1977; Schwarz, 1978). BIC evaluates the goodness of fit of the model while scaling with the sample size or the number of variables, whereas AIC does not take into account the sample size or the number of variables. The model with the most explanatory power has the highest r^2 value and lowest AIC and BIC (i.e., the greatest percentage of variation explained per degree of freedom).

I performed two separate sets of GLM analyses, one for the lecture-only course and one for the lab course. If any variable were determined to have an effect on the postcourse inventory scores in either of the two populations (lecture only or lab), that variable would help explain differences between the lecture-only and the lab courses.

RESULTS

From the pre- to the postcourse inventory, mean scores increased 12.9% for students in the lecture-only course and 17.3% for students in the lab course. That is a 33% greater increase for the lab course compared with the lecture-only course (Fig. 1 and Table IV).

There was no significant difference found between the standardized inventory scores of the 28 GCI questions and those of all 37 inventory questions (Fig. 2). This result was consistent for both the pre- and the postcourse inventories.

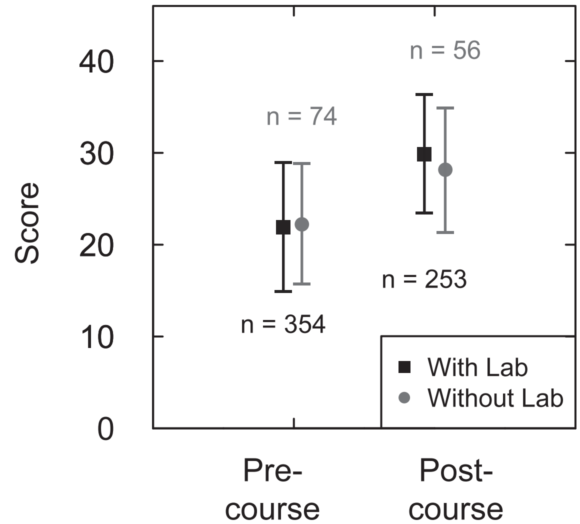

FIGURE 1: Mean ± 1 SD of assessment scores for analysis including all scores. The four numbers represent the four sample sizes for each of the four populations.

Even though the nine questions created for this inventory were not tested for reliability and validity, they had no effect on student performance. Therefore, I focus on the results of all 37 concept inventory questions for the discussion.

Within the analysis of the inventory scores for the matched-identification numbers, students demonstrated an increase of 11.9% for the lecture-only course and 17.0% for the lab course (Fig. 3 and Table IV). That is a 43% greater increase for the lab course compared with the lecture-only course.

The null model of randomly selected student responses produced a mean inventory score of 9.1 with a standard deviation (SD) of 2.4; this translates to a mean score of 25% if out of 37 total questions and 20% if out of 46 total correct responses. Because all inventory scores (pre- and postcourse) were greater than 47.6%, there is a low likelihood that the mean scores (pre- or postcourse for both courses) were influenced by random guessing.

There was a significant difference between the pre- and the postcourse inventory scores for both courses ($p < 0.001$; Table V). The postcourse inventory scores were significantly

TABLE IV: Absolute scores (and percentage scores) for the student concept inventory.

	Precourse	Postcourse
Complete analysis		
With lab	21.9 (47.6%)	29.9 (64.9%)
Lecture only	22.2 (48.3%)	28.2 (61.2%)
Matched identification		
With lab	22.8 (49.5%)	30.6 (66.6%)
Lecture only	22.6 (49.2%)	28.1 (61.1%)
12 lab-specific questions		
With lab	7.8 (49%)	11.2 (70%)
Lecture only	8.5 (53%)	10.3 (64%)

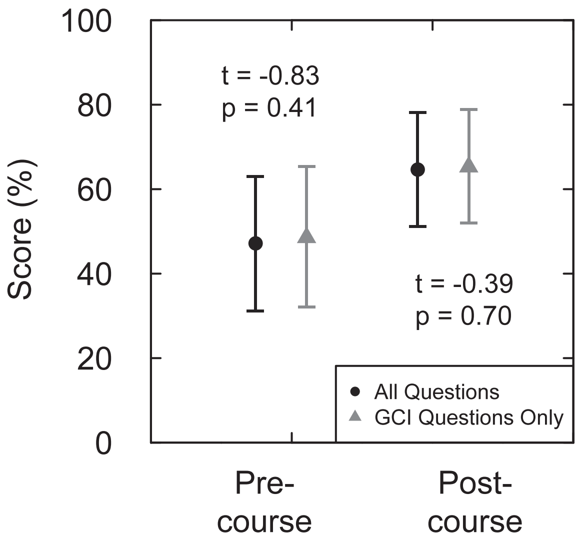

FIGURE 2: Standardized scores for pre- and postcourse concept inventories consisting of the 28 GCI questions only and all 37 questions. The results for the two *t*-tests are also presented: one between the two precourse score populations and one between the two postcourse score populations.

greater for both courses. Using a paired-sample *t*-test, a significant difference was found between the pre- and the postcourse matched-identification scores for both courses ($p < 0.001$; Table VI).

Within the analysis of the 12 questions that were specific to exercises performed in the laboratory section of the lab course, there was an increase of 21% for the lab course and 11% for the lecture-only course (Table IV). That is an 84% greater increase for the lab course compared with the lecture-only course (Fig. 4 and Table VII).

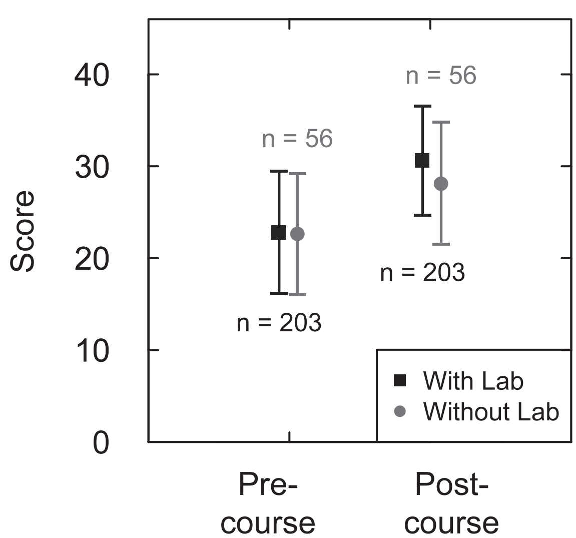

FIGURE 3: Mean ± 1 SD of assessment scores for analysis of the matched-identification scores. The four numbers represent the four sample sizes for each of the four populations.

TABLE V: *T*-test results for comparisons of all assessment scores. "Y" or "N" denotes whether the test was statistically significant at $p < 0.05$.

With Lab Pre- vs. Postcourse	Lecture Only Pre- vs. Postcourse
$p < 0.001$ (Y)	$p < 0.001$ (Y)
$t = -14.28$	$t = -4.96$
$n = 354$ (pre)	$n = 74$ (pre)
$n = 253$ (post)	$n = 56$ (post)

There were no effects from instructor, class standing, or academic major on either of the GLMs (Fig. 5 and Table VIII). The GLM for both the lab course and the lecture-only course had minimal increases in r^2 statistics and the lowest AIC and BIC results from the univariate models consisting of only the inventory scores without other variables.

DISCUSSION

Students in both courses (lecture-only and lecture plus lab) demonstrated a significant increase in concept inventory scores over the span of each course (Figs. 1 and 2 and Tables IV–VII). However, the inventory scores for the lab course increased 33% more than those for lecture-only course and 43% more within the matched-identification score analysis. Because the same two instructors teach the same material in the lecture sections of both courses, the greater increase of scores for students enrolled in the lab course reflects that those students gained more conceptual knowledge in the course with an associated laboratory.

Because of the experimental design, a disadvantage of this study is that I cannot directly distinguish between the lack of a lab and the shortened length of the course (3 weeks as opposed to 13 weeks) as the cause for the lower learning gains in the lecture-only course. I was unable to manipulate either course to setup a proper control for the experiment. It would be unethical, and possibly detrimental to the students, to remove the laboratory section from the lab course. However, there are two novel and important factors of this study that overcome this one disadvantage. (1) The same two instructors teach the same material in the lecture section of both the lab course and the lecture-only course. Introductory courses often either have a required lab or do not. It is less common to have the same instructors teach two versions of a course, particularly without adding material to the lecture to compensate for the lack of a laboratory component. (2) An independent instrument was employed, and the same concept inventory questions were given to all

TABLE VI: Paired-sample *t*-test results for comparisons of matched-identification assessment scores. "Y" or "N" denotes whether the test was statistically significant at $p < 0.05$.

With Lab Pre- vs. Postcourse	Lecture Only Pre- vs. Postcourse
$p < 0.001$ (Y)	$p < 0.001$ (Y)
$t = -21.89$	$t = -7.58$
$n = 203$ (pre)	$n = 56$ (pre)
$n = 203$ (post)	$n = 56$ (post)

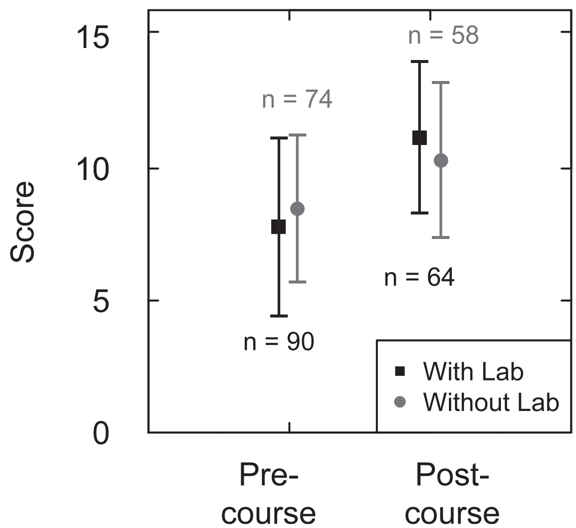

FIGURE 4: Mean ± 1 SD of concept inventory scores for analysis of the 12 questions (totaling 16 points) that were directly asked in one of the questions within a lab exercise. The four numbers represent the four sample sizes for each of the four populations.

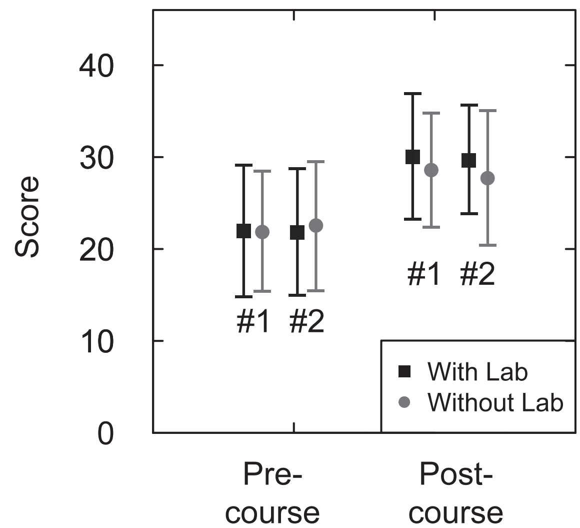

FIGURE 5: Mean ± 1 SD of assessment scores for analysis including all scores separated by the two instructors. The "#1" represents scores from instructor 1, and the "#2" represents scores from instructor 2.

students pre- and postcourse. By using the GCI, I am able to conduct a controlled inventory using a reliability- and validity-tested instrument to gauge student conceptual understanding prior to and after the course. This method provides a measure independent from the graded evaluations administered by the instructors. Furthermore, providing the same questions pre- and postcourse is the most direct means of assessing student conceptual knowledge outcomes. If the questions change or if some other measure is used, there is a chance for error in gauging what the students have gained.

Even though I was unable to control for course length, previous literature indicates that intensive courses are as effective as full-length courses in that these courses produce the same student learning gains (Daniel, 2000, and references within; Davies, 2006, and references within; Kucsera and Zimmaro, 2010; Nasiri and Shokrpour, 2012). Inventory scores for time-shortened and intensive courses are consistently the same as for traditional-length courses in a range of disciplines, including the Earth sciences (Waechter, 1966). This evidence suggests that the shortened nature of the lecture-only course did not lead to the lower learning gains compared with the lab course.

If the shortened length of the lecture-only course was not the cause for the poorer learning gains observed, the lack of a laboratory sections was the most likely cause. There has been little research into the effects of laboratory components accompanying university science courses (Hofstein and Lunetta, 1982, 2004). Specifically for the Earth sciences, only one study has examined the importance of having a laboratory section along with a lecture section (Neilson et al., 2010). Neilson et al. found that students enrolled in an optional laboratory section of an introductory Earth science course performed better than students not enrolled in the laboratory section; the students obtained a greater amount of total course points. The results of the present study corroborate the findings of Neilson et al. and support the long-held assumption that laboratory sections increase student learning outcomes (Hofstein and Lunetta, 2004). Thus, it is likely that the 33% greater increase in student concept knowledge in the lab course compared to the lecture-only course was a direct result of the laboratory component. This is an important affirmation of the effectiveness of university Earth science laboratory sections. To further support the importance of Earth science laboratories in university student education, researchers need to conduct additional quantitative and qualitative case studies into the outcomes of introductory course laboratories. In addition, assessment of the effectiveness of upper-level Earth science courses is needed to reveal how laboratories aid student learning in more advanced geoscience courses.

In this study, the course that contained a lecture and a lab provided the students with approximately twice the amount of class time as the lecture-only course. Therefore,

TABLE VII: *t*-test results for comparisons of concept inventory scores for the 12 questions (totaling 16 points) that were explicitly asked in one of the questions within a lab exercise. "Y" or "N" denotes whether the test was statistically significant at $p < 0.05$.

With Lab Pre- vs. Postcourse	Lecture Only Pre- vs. Postcourse	With Lab vs. Lecture Only Postcourse
$p < 0.001$ (Y)	$p < 0.001$ (Y)	$p = 0.11$ (N)
$t = -6.66$	$t = -3.60$	$t = -1.61$
$n = 90$ (pre)	$n = 74$ (pre)	$n = 64$ (lab)
$n = 64$ (post)	$n = 58$ (post)	$n = 58$ (no lab)

TABLE VIII: GLM results taking into account the different instructors on the assessment scores.

	p value	r^2	BIC	AIC
With lab				
Inventory only	<0.001	0.41	807.13	798.44
Inventory + different instructors	<0.001	0.41	811.02	799.43
Inventory + class standing (year)	<0.001	0.45	821.28	798.10
Inventory + major	<0.001	0.46	810.07	792.67
Inventory + class standing (year) + major	<0.001	0.46	827.95	798.97
Lecture only				
Inventory only	<0.001	0.44	344.47	338.45
Inventory + different instructors	<0.001	0.47	349.93	339.89
Inventory + class standing (year)	<0.001	0.47	358.40	344.35
Inventory + major	<0.001	0.45	351.92	341.89
Inventory + class standing (year) + major	<0.001	0.47	366.02	347.95

the students in the lab course could have gained a greater amount of conceptual knowledge because they had a greater amount of time on task. I tested this possibility by examining the inventory scores of a subset of 12 of the total 37 questions (Fig. 4 and Table VII). These 12 questions contained content that was within a question the students had to answer for at least one of the lab exercises. Students in the lab course had a greater opportunity to actively engage in the material covered by these 12 questions. There was only a marginally significant difference between the post-course score populations of these 12 questions ($p = 0.11$). Although this is greater than the often employed threshold of $p = 0.05$, I contend this marginally significant difference provides evidence that the students in the lab course are gaining a greater amount of conceptual knowledge on this subset of inventory questions. Furthermore, the students in the lab course had an increased score 10% greater than that of students in the lecture-only course (Table IV). Thus, the greater increase in concept knowledge on these 12 questions by students in the lab course is evidence the lab course directly aids in student learning outcomes (Fig. 4 and Table VII).

Some may argue the greater increase of scores on these 12 questions is not sufficient evidence (and not statistically great enough) that the lab content was leading to the greater increase in student conceptual knowledge. In this case, whether the greater increase in inventory scores on these 12 questions was due to greater time on task, reinforcement of the lecture material, or students learning the material for the first time, the students in the lab section still had greater increases in conceptual knowledge (on these 12 questions and all 37 questions). The laboratory section provided students with a worthy complement to the lecture section, leading to a greater understanding of Earth science concepts.

As is the case with many time-shortened or intensive courses (Daniel, 2000), the students in the lecture-only course were generally further along in their university careers than were the students in the course with the lab (Table III). Students further into their university career often have a stronger work ethic (Daniel, 2000; Davies, 2006). If this factor did affect student performance on the concept inventory, I predicted it would lead to higher scores. Given that the postcourse scores for the lecture-only course were lower than those for the lab course, it can reasonably be concluded that the greater number of third- and fourth-year students was not a factor. Conversely, if the older and nontraditional students were taking the intensive, lecture-only course because of poor planning early in their university career, lack of desire to take a science course, or some other negative motive, I would not expect their scores to be higher. However, the mean precourse scores (and assessment score SD) for both courses were approximately the same (Figs. 1–3 and Table IV). In addition, there was no effect of student academic year on the increase in pre- to postinventory scores as determined by the two separate GLMs (Table VIII). To further examine this complex issue, I plan to conduct future examinations where student academic year is controlled.

The length of lecture period differed between the two courses. Class length for the lab course was 50 min three times per week. Class length for the lecture-only course was 3 h per day (with two 10-min breaks), 5 d per week. The longer length of lecture period for the lecture-only course may have resulted in lower learning gains. It is difficult for students to continually take in and process information for more than 2.5 h (Hartley and Davies, 1978; Wankat, 2002). However, a lengthy lecture period is often a factor in time-shortened, intensive courses. If the long lecture period led to lower learning gains, this should be a consistent finding for all time-shortened, intensive courses. Because I established that this is not the case (as explained earlier), the length of the lecture periods most likely had little impact on the difference in learning gains between the two courses. Furthermore, research has demonstrated the student learning outcomes are not affected by longer, intensive lecture periods (Daniel, 2000).

There was no statistical evidence of any difference in inventory scores (pre- or postcourse) among the two instructors, 4 y, or academic majors (Fig. 5). None of these variables led to a greater increase in scores from pre- to postcourse (Table VIII). This is further evidence that the differences between the two courses were caused by the inclusion or exclusion of the laboratory component.

An alternative qualitative interpretation of the results from this study is that the increase in inventory scores in the lab course is similar to the increase for the lecture-only course. This interpretation suggests that students in the lab course did not gain more conceptual knowledge than students in the lecture-only course. When the time and resources put into the lab course are weighed, the increase in conceptual knowledge gains may not be resource beneficial. Two responses to this interpretation are (1) that any increase in conceptual knowledge is worth the cost. To better our society, people need to understand how the world works to be informed decision makers. Particularly, at the university level, this may be the students' one science course, specifically preparing preservice teachers, whose science preparation is often lacking (McDermott, 1990; Supovitz and Turner, 2000; Loverude et al., 2011). Without laboratories for experiential learning, these students may be unprepared to lead laboratory activities for their own future students. Therefore, providing all students at all levels with the best opportunity for gaining knowledge is worth the cost. (2) This study is a first step in examining the effectiveness of laboratories compared with lecture-only courses. Additional studies using the GCI will further explore student learning gains in courses with and without laboratories. I will employ additional sample populations, plus a specialized assessment instrument geared specifically toward measuring the effects of laboratory courses. These data will help to determine whether the greater increase in conceptual knowledge by students in the lab course is a consistent result, whether a reassessment of the laboratory activities may be in order, and whether the mechanisms by which laboratories increase student understanding match the hypotheses developed from anecdotal observations.

CONCLUSIONS

To conduct a direct test of the effect of a laboratory section on student learning gains, I administered the same pre- and postcourse concept inventory to students in two courses, one that contained a laboratory section one that was only a lecture. The same two instructors taught the same material in both courses. The mean increase in inventory score for the course with a laboratory was 33% greater (43% greater for matched-identification score analysis) than the mean score increase for the lecture-only course. The only notable difference between the courses was that the lecture-only course was a time-shortened course. Because overwhelming research exists demonstrating that time-shortened, intensive university courses produce the same student learning gains as traditional-length courses, the inclusion of the laboratory section most likely led to the greater increase in student conceptual knowledge in that course. This result demonstrates the importance of a laboratory section accompanying the lecture section of introductory-level, university Earth science courses to maximize the educational outcomes for students.

Acknowledgments

Thank you to Dr. Kristen St. John and two anonymous reviewers for thoughtful and helpful comments that greatly improved this manuscript. I thank the Centre for Teaching and Learning at the University of Alberta, and in particular Academic Director Dr. Heather Kanuka, for enthusiastically supporting this project, both with insights into the issues and with funding to allow me to do this project. I also thank Lindsey Leighton and Bob Luth for supporting me throughout the span of this project and for their helpful comments leading to a better project. I thank Emily Stafford, Julie Libarkin, and the Michigan State University Geocognition Research Laboratory for help with the manuscript. A special thank you goes to instructors Fred Clark, Solweig Balzer, and David Chesterman for volunteering to take part in this study.

REFERENCES

Akaike, H. 1977. On entropy maximization principle. *In* Krishnaiah, P.R., ed., Applications of statistics. Amsterdam, the Netherlands: North-Holland Publishing, p. 27–41.

Baird, J.R. 1990. Metacognition, purposeful enquiry and conceptual change. *In* Hegarty-Hazel, E., ed., The student laboratory and the science curriculum. London: Routledge, p. 183–200.

Blosser, P. 1983. The role of laboratory in science education. *School Science and Mathematics,* 83:165–169.

Cowan, N. 1998. Visual and auditory working memory capacity. *Trends in Cognitive Science,* 2:77–78.

Daniel, E.L. 2000. A review of time-shortened courses across disciplines. *College Student Journal,* 34:298–308.

Davies, W.M. 2006. Intensive teaching formats: A review. *Issues in Education Research,* 16:1–20.

Doran, R.L. 1978. Assessing the outcomes of science laboratory activities. *Science Education,* 62:401–409.

Gadt-Johnson, C.D., and Price, G.E. 2000. Comparing students with high and low preferences for tactile learning. *Education,* 120:581–585.

Hartley, J., and Davies, I. 1978. Note taking: A critical review. *Programmed Learning and Education Technology,* 15:207–224.

Hofstein, A., and Lunetta, V.N. 1982. The role of laboratory in science teaching: Neglected aspects of research. *Review of Education Research,* 52:201–217.

Hofstein, A., and Lunetta, V.N. 2004. The laboratory in science education: Foundations for the twenty-first century. *Science Education,* 88:28–54.

Huysken, K., Neilson, K.G., and Kilibarda, Z. 2011. Assessing the effect of laboratory enrollment on lecture performance in introductory Earth science courses. *Geological Society of America Abstracts with Programs,* 43:146.

Jolley, A., Lane, E., Kennedy, B., and Frappe-Seneclauze, T. 2012. SPESS: A new instrument for measuring student perceptions in Earth and ocean science. *Journal of Geoscience Education,* 60:83–91.

Kucsera, J.V., and Zimmaro, D.M. 2010. Comparing the effectiveness of intensive and traditional courses. *College Teaching,* 58:62–68.

Libarkin, J.C., and Anderson, S.W. 2005. Assessment of learning in entry-level geoscience courses: Results from the Geoscience Cognitive Inventory. *Journal of Geoscience Education,* 54:394–401.

Libarkin, J.C., and Anderson, S.W. 2006. The Geoscience Concept Inventory: Application of Rasch analysis to concept inventory development in higher education. *In* Liu, X., and Boone, W., eds., Applications of Rasch measurement in science education. Maple Grove, MN: JAM Publishers, p. 45–73.

Libarkin, J.C., and Kurdzel, J.P. 2001. Research methodologies for science education: Strategies for productive assessment. *Journal of Geoscience Education,* 49:300–304.

Loverude, M.E., Gonzelas, B.L., and Nanes, R. 2011. Inquiry-based course in physics and chemistry for preservice K–8 teachers. *Physical Review Special Topics—Physics Education Research,* 7:1–18.

Matz, R.L., Rothman, E.D., Krajcik, J.S., and Banaszak-Holl, M.M. 2012. Concurrent enrollment in lecture and laboratory enhances student performance and retention. *Journal of Research in Science Teaching,* 49:659–682.

McDermott, L.C. 1990. A perspective on teacher preparation in physics and other sciences: The need for special science courses for teachers. *American Journal of Physics,* 58:734–742.

Nasiri, E., and Shokrpour, N. 2012. Comparison of intensive and non-intensive English courses and their effects on the student's performance in an EFL university context. *European Scientific Journal,* 8:127–137.

Neilson, K.G., Huysken, K., and Kilibarda, Z. 2010. Assessing the impact of geoscience laboratories on student learning: Who benefits from introductory labs? *Journal of Geoscience Education,* 58:43–50.

Penney, C. 1989. Modality effects of short-term and the structure verbal memory. *Memory and Cognition,* 17:398–422.

Perkins, D. 2004. Scholarship of teaching and learning, assessment, and the *Journal of Geoscience Education*. *Journal of Geoscience Education,* 52:113–114.

Prades, A., and Espinar, S.R. 2010. Laboratory assessment in chemistry: An analysis of the adequacy of the assessment process. *Assessment and Evaluation in Higher Education,* 35:449–461.

Prain, V., and Waldrip, B. 2006. An explanatory study of teachers' and students' use of multi-modal representations of concepts in primary science. *International Journal of Science Education,* 28:1843–1866.

Schwarz, G.E. 1978. Estimating the dimension of a model. *Annals of Statistics,* 6:461–464.

Supovitz, J.A., and Turner, H.M. 2000. The effects of professional development on science teacher practices and classroom culture. *Journal of Research in Science Teaching,* 37:963–980.

Waechter, R.F. 1966. A comparison of achievement and retention by college junior students in an Earth science course after learning under massed and spaced conditions [Ph.D. dissertation]. State College, PA: Pennsylvania State University. p. 1–139.

Wankat, P. 2002. The effective efficient professor: Teaching, scholarship, and service. Boston, MA: Allyn and Bacon.