The Impacts of Curriculum Reform for an Introductory Undergraduate Biology Course Sequence

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Anna Jo Juanita Auerbach
May 2016

Copyright © 2016 by Anna Jo Auerbach
All rights reserved.

DEDICATION

“Tell me and I forget, teach me and I may remember, involve me and I learn.”

~Benjamin Franklin

This work is dedicated to all the learners of the world and the lessons that guide them.

ACKNOWLEDGEMENTS

I consider myself extremely fortunate to have so many people that deserve an acknowledgement for their contribution to my work. First and foremost, I would like to thank my advisor Dr. Elisabeth Schussler, for spearheading a curriculum reform guided by *Vision and Change* recommendations*,* taking a chance on a non-major, and for countless hours of advisement, feedback, and mentoring in areas related to my research. I would also like to thank my committee comprised of Drs. Randall Small, Jennifer Ann Morrow, and Jennifer Schweitzer for their valuable feedback, suggestions, and questions. Thank you to the Schussler lab both past and present: Sarah Dalrymple, Ben England, Miranda Chen, and Jennifer Brigati. Thank you to Stephanie Robinson for her time as an inter-relator observer. I owe thanks to all of the graduate teaching assistants of the reformed courses for their help with data collection, the undergraduate students and faculty members for their participation in this work, and the department of Ecology and Evolutionary Biology for supporting such research endeavors. This work would not have been possible without all of you.

I am beyond grateful for my core support system that is my husband Jeremy Auerbach, my mother Susan Davis, and my best friend Nicole Fogarty. Without their belief in me, patience, and support of me as a person I would not be where I am today. I love you guys! I must also give thanks for the music of the Talking Heads, Tool, and Pink Floyd, without which I may not have made it through months of data analysis and writing. Last, but not least, I would like to give thanks for the practice of meditation. I learned far greater than the sum of this research throughout my experience and owe much of what I gained to daily meditation.

ABSTRACT

For decades, educators and policy makers have called for reform in undergraduate biology education to produce a workforce capable of navigating the challenges of society today (NSB, 1986). The American Association for the Advancement of Science (AAAS) outlined recommendations to enact these changes in the *Vision and Change in Undergraduate Biology Education* final report (AAAS, 2011). These recommendations included restructuring curricula to focus on core biological concepts and competencies, implemented via student-centered practices, with a campus-wide commitment to change, and involving the entire biology community in the reform. The University of Tennessee, Division of Biology, implemented curriculum changes to their majors' introductory biology sequence based on these recommendations in 2012. This curriculum reform had several goals: integrate the concepts and competencies into the two-course sequence, use more student-centered instructional approaches, and create discussion sections that targeted process of science skills. The goal of this dissertation was to document instructional changes throughout reform, how faculty perceptions of instruction changed, and what impact the reform had on student outcomes such as scientific literacy skills. The first chapter documents the baseline use of student-centered approaches, or active learning, during the first year of curriculum reform. The second chapter built on the first by conducting three years of classroom observations and faculty interviews to document the perception and use of student-centered instructional practices by core faculty of the introductory courses. Chapter three measures student scientific literacy gain and retention the year prior to and the first year of the curriculum reform to compare the learning outcomes. Overall, the instructors

increased the use of active learning in their classes over time and changed their perceptions of how students learn. Students who experienced the new discussion curricula had greater gains in scientific literacy, and retained more information about research design, despite a loss of lab time. A reform specifically designed to target instructor practices and process of science skills succeeded. Collectively, these studies provide evidence for the importance of collaboration, reflection, and time for instructors to achieve the pedagogical changes needed for modern undergraduate biology reform.

TABLE OF CONTENTS

INTRODUCTION **1**

CHAPTER I INSTRUCTOR USE OF GROUP ACTIVE LEARNING IN AN INTRODUCTORY BIOLOGY SEQUENCE **6**

Abstract *7*

Introduction *8*

Methods *12*

Results *17*

Discussion *22*

CHAPTER II A VISION AND CHANGE REFORM OF INTRODUCTORY BIOLOGY IMPROVES FACULTY PERCEPTIONS AND USE OF ACTIVE LEARNING **27**

Abstract *28*

Introduction *28*

Methods *32*

Results *44*

Discussion *64*

Appendix *69*

CHAPTER III CURRICULUM ALIGNMENT WITH VISION AND CHANGE IMPROVES STUDENT SCIENTIFIC LITERACY **70**

Abstract *71*

Introduction *71*

Methods *76*

Results *83*

Discussion *91*

Appendix *96*

CONCLUSION **97**

REFERENCES **100**

VITA **110**

LIST OF TABLES

Table 1. Classification of classroom event categories 14

Table 2. Average proportion of time spent in classroom events by instructor. 18

Table 3. Summary of instructor demographics. 33

Table 4. Average proportion of class time in active learning across reform. 46

Table 5. Summary of interview categories and themes. 49

Table 6. Schedule of three TOSLS implementations for each cohort. 78

LIST OF FIGURES

Figure 1. Average event type by instructor 20

Figure 2. Average percent class time in classroom event by course 21

Figure 3. Proportion of class time in active learning and group active learning across three years of reform 45

Figure 4. Overall gain from pre-test to post-test in scientific literacy for cohort 1 and cohort 2. 85

Figure 5. Gain from pre-test to post-test in the skill of identify a valid argument for cohort 1 and cohort 2. 86

Figure 6. Gain from pre-test to post-test in the skill of research design for cohort 1 and cohort 2. 88

Figure 7. Pre, post, and retention test scores for cohort 1 and cohort 2 89

Figure 8. Retention from post-test to retention measure in the skill of identify research design for cohort 1 and cohort 2. 90

INTRODUCTION

For decades, educators and policy makers have called for reform in undergraduate biology education to produce a workforce capable of navigating the challenges of society today (NSB, 1986). National efforts to reform undergraduate science instruction, particularly introductory "gateway" courses for science majors, have been underway for over 15 years with limited success (AAAS, 2011; Handelsman et al., 2004; NRC, 2004; NSF, 1996). As a result, the American Association for the Advancement of Science (AAAS) began conversations with stakeholders across the country with the goal of creating a framework for introductory biology that would meet the demands of undergraduate biology education and be broadly adaptable to individual programs and institutions (AAAS, 2009). These conversations led to a national conference and in 2011 recommendations for restructuring undergraduate biology were published in the *Vision and Change in Undergraduate Biology Education* final conference report (AAAS, 2011). The recommendations included restructuring curricula to focus on core biological concepts and competencies, implemented via student-centered practices, with a campus-wide commitment to change, and involving the entire biology community in the reform. The core concepts represent overarching principles that permeate all fields of biology. However, the successful practice of biology requires more than just an understanding of biological concepts; it requires an ability to apply skills, or competencies, such as being able to make decisions using scientific knowledge and apply the process of science.

However, merely changing what is taught in the classroom is not enough to enact the reforms needed for the students of today. There is also a need to change *how* we

teach. Student-centered pedagogies, also known as active learning, are approaches that engage student thinking through the use of class activities that require students to reflect upon their ideas and their application (Collins & O'Brien, 2003). Active learning in introductory biology courses has been found to improve student performance, including the success of underrepresented students in the courses (Armbruster, Patel, Johnson, & Weiss, 2009; Freeman, et al*.,* 2007; Michael, 2006), as well as reduce the achievement gap between non-disadvantaged and disadvantaged learners (Haak et al., 2011). Active learning has also been shown to lower failure rates in undergraduate biology courses (Freeman, Haak, & Wenderoth, 2011). A multi-disciplinary meta-analysis provided robust evidence that active learning increases student performance regardless of science discipline (Freeman et al., 2014).

What is also clear, however, is that how active learning strategies are implemented makes a difference. Peer discussion has been shown to benefit student learning across all levels and disciplines, including undergraduate biology students (Smith et al., 2009; Smith et al., 2011). Webb (1989) found that student achievement was increased after students gave explanations within small groups. Coleman, Brown, and Rivkin (1997) demonstrated that biology students teaching each other experienced stronger learning effects compared to those just summarizing. Classrooms that do not allow for student collaboration and discussion may impede students in their learning by not facilitating opportunities for students to test their own ideas, interact with information via numerous sensory contexts, and synthesize those ideas with others.

The University of Tennessee, Knoxville, where this dissertation research was conducted, recently implemented a reform of its introductory biology sequence guided

by the *Vision and Change in Undergraduate Biology Education* report (AAAS, 2011). The faculty who taught the two courses formed a curriculum reform community, funded by a grant from the National Science Foundation (NSF DUE 1245215; PI Schussler) to plan for and enact changes to the concepts, competencies, and active learning pedagogies employed in the courses. One way to enact the changes is through a top-down process where decisions are dictated to faculty. Another approach is to work together as an instructional community to make decisions about curriculum changes while still allowing individuals to have some autonomy to match their individual teaching styles. A learning community paired with some flexibility for individual faculty change was preferred over the creation and implementation of common course modules because the literature suggests that grassroots change is superior to top-down initiatives at universities (Henderson, Beach, & Finkelstein, 2011; Silverthorn, Thorn, & Svinicki, 2006).

The two courses that made up the original majors' sequence were 4-credit hour courses that were each paired with an associated lab. One course focused on biodiversity, ecology, and evolution, while the other focused on cellular and molecular content. The curriculum reform not only revised the core concepts and competencies used within each course, but also changed the structure of the course delivery. The labs from each semester course were combined into a single semester to create one 2-credit hour lab course with discussion. Weekly small group discussion sections were also added to each 3-credit hour lecture class, resulting in a loss of one third of lecture time compared to the previous year. Thus, both the old and the new sequence were 8 credit

hours and retained the same topical foci, but the reformed sequence had less lab time and more discussion time then the previous version of the sequence.

The discussion curricula were designed in a community context that engaged a volunteer subset of the entire biology community including undergraduate and graduate students, post docs, and faculty. These groups (one for each discussion) took the academic year of 2013-2014 to create the discussion syllabi, assessment approach, module and weekly topics, learning objectives, and some of the lesson plans. The learning objectives for each discussion were aligned with the *Vision and Change* core competencies and intended to foster scientific literacy by using primary literature as the focus of active learning activities and student collaboration.

The curriculum reform took place over three years. During the first year (2012 – 2013), a few faculty piloted the new concepts and competencies within their courses. The next year (2013 – 2014), all faculty members used the new concepts and competencies within their courses. In the third year (2014 – 2015), the structural changes were implemented and the new discussion sections were added to each course. The goal of this dissertation was to document instructional changes throughout reform, how faculty perceptions of instruction changed, and what impact the reform had on student scientific literacy skills.

The first chapter documents the baseline use of individual active learning and group active learning (students using peer discussion) pedagogies by faculty teaching courses in the introductory biology sequence during the first year of curriculum reform (2012 – 2013). The intent of this study was to collect data about current classroom practices, which would then be used to provide feedback to faculty and guide

professional development needs in the context of the curriculum reform. Classroom observations documented the occurrence and duration of classroom events (including lecture, clicker questions, group discussion) over one academic year. The specific research objectives were to document the occurrence of and time spent on different classroom events, compare faculty implementation of individual and group active learning in their classrooms, and examine the overall implementation of individual and group active learning across the sequence of courses.

The second chapter built on the first by conducting classroom observations and faculty interviews across the three years of reform to document the perception and use of student-centered instructional practices by core faculty of the introductory courses. The research questions guiding the research of this chapter were to determine how instructors' active learning practices and active learning beliefs were changing throughout the process of reform.

The third chapter looks at the reform from the standpoint of student outcomes. The specific research questions were to determine if curriculum alignment with *Vision and Change* competencies resulted in a higher gain or retention of student scientific literacy skills. Student scientific literacy test gains were compared in a pre-post fashion between two cohorts: one group that experienced the previous curricula without the discussions and one group that experienced the reformed curricula with the discussions. Cohort scores on the same scientific literacy test were also compared one semester (4 months) after completion of the introductory sequence to determine if there were differences in scientific literacy retention between the groups.

CHAPTER I

INSTRUCTOR USE OF GROUP ACTIVE LEARNING IN AN INTRODUCTORY BIOLOGY SEQUENCE

A version of this chapter was originally published by Anna Jo Auerbach and Elisabeth E. Schussler:

Anna Jo Auerbach, Elisabeth E. Schussler. (in press). “Instructor Use of Group Active Learning in an Introductory Biology Sequence.” *Journal of College Science Teaching.*

This above published article was revised with minor text edits in order to be included in this dissertation. The first author contributed to the research concept, data collection, data analysis, and manuscript writing. The second author contributed to the research concept, the inter-rater reliability of observations, and manuscript writing.

Abstract

Active learning (or learner-centered) pedagogies have been shown to enhance student learning in introductory biology courses. Student collaboration has also been shown to enhance student learning and may be a critical part of effective active learning practices. This study focused on documenting the use of individual active learning and group active learning pedagogies of faculty teaching the two course introductory biology sequence at one large university. Monthly observations recorded the frequency, length, and type of active learning strategies (e.g., clicker questions, verbal questions, and activities) used in classes. Differences were found in the use of active learning type as well as use of individual and group active learning when comparing instructors independently and by course. These observational data were used as a tool to provide feedback to the instructors and inform future professional development. These findings suggest that professional development for biology instructors should focus on the

planning for and implementation of group active learning strategies in introductory courses.

Introduction

National efforts to reform undergraduate science instruction, particularly introductory, gateway courses for science majors, have been underway for over 15 years with limited success (AAAS, 2011; Handelsman et al., 2004; NRC, 2004; NSF, 1996). In particular, faculty have been challenged to integrate more active learning pedagogies into their courses, where students are encouraged to construct an understanding of the material rather than passively receiving information from a lecture (NRC, 2000). Active learning pedagogies can vary widely, from instructor-generated verbal questions that students think about or respond to on their own, to small group activities and discussions, to personal response "clicker" questions, with or without group discussion. In each case, the goal is to help students take an active role in their own learning while in class.

Consistent with the assumption that active learning increases student learning, courses focused on these student-centered pedagogies have been shown to increase student academic performance (Armbruster et al., 2009; Freeman et al., 2007; Hake, 1998; Knight & Wood, 2005), as well as reduce the achievement gap between non-disadvantaged and disadvantaged learners (Haak et al., 2011). Active learning has also been shown to lower failure rates in undergraduate biology courses (Freeman, Haak, & Wenderoth, 2011). Derting and Ebert-May (2010) reported that students who participated in introductory coursework with active learning had a better understanding of science as a process by their senior year. A recent multi-disciplinary meta-analysis

provided robust evidence that active learning increases student performance regardless of discipline (Freeman et al., 2014).

While students acknowledge that active learning increases engagement in classes (Welsh, 2012), they also articulate that some instructors are better at using active learning pedagogies than others. Indeed, how active learning pedagogies are employed make a difference in student learning gains. Smith et al. (2009) and Smith et al. (2011) documented gains in student learning from clicker questions and found that peer discussion in combination with instructor explanation was critical to maximizing student learning using this pedagogy. Turpen and Finkelstein (2010) found that the nature of classroom discussion in general appears to be important in helping physics students make sense of the material, yet how discussion is employed in the classroom is extremely variable from professor to professor, with some fostering much higher levels of student discussion than others (Turpen & Finkelstein, 2009). This may be why some active learning reforms do not result in the same learning gains as others; how the instructor implements active learning may make a profound difference to its success (Andrews et al., 2011).

Student talk, specifically self-explanation, plays a vital role in the learning process as shown in both psychological (Chi et al*.,* 1994; Webb, 1989) and educational research literature (Coleman, Brown & Rivkin, 1997; Smith et al*.,* 2009). Chi et al*.* (1994) proposed that self-explanation was the cognitive process responsible for enabling the incorporation of new knowledge into existing knowledge. Webb (1989) found that student achievement was increased after students gave explanations within small groups. Coleman, Brown, and Rivkin (1997) demonstrated that biology students

participating in teaching each other experienced stronger learning effects compared to those participating in summarizing. Student talk, used in such group active learning pedagogies, has been shown to benefit students from all levels and disciplines, including undergraduate biology students (Smith et al., 2009; Smith et al., 2011). Indeed, cognitive studies have found that multiple sensory inputs, such as both seeing and hearing the same information, has been shown to enhance memory, and practice increases fluency with the material (Herman, Munk, & Engel, 2004; Shaw, McGaugh, & Rose, 1990); both sensory inputs and practice would be anticipated to increase in classrooms where students talk to each other about the material they are learning. Classrooms that do not allow for student collaboration and discussion may impede students in their learning by not facilitating opportunities for students to test their own ideas, interact with information via numerous sensory contexts, and synthesize those ideas with others.

This study focused on documenting the use of individual active learning and group active learning pedagogies by faculty teaching courses in the introductory biology sequence at one large university. The institution where the study was conducted is in the process of reforming its introductory biology sequence guided by the *Vision and Change in Undergraduate Biology Education* report (AAAS, 2011). The faculty who teach the courses formed a curriculum reform community in 2012-2013, funded by a grant from the National Science Foundation, to plan for and enact changes to the concepts, competencies, and active learning pedagogies employed in the introductory majors' course sequence. As part of this process, the program undertook a baseline

study of current active learning practices in the courses comprising the introductory sequence.

Introductory biology courses at large universities are often taught in multiple lecture sections by different instructors. Smith et al. (2014) recently conducted observations in a range of STEM courses at a large university and reported on the variety of classroom practices employed by different professors. These differences in active learning practices could lead to different student classroom experiences and potentially learning outcomes, particularly depending on whether the active learning was enacted with group discussion or not. Smith et al. (2014) also advocated for the use of observational data as a guide for faculty professional development. Therefore, the intent of this study was to collect data about current classroom practices in multiple sections of an introductory biology sequence at one university, which would then be used to provide feedback for faculty and guide professional development needs in the context of the curriculum reform. Classroom observations documented the occurrence and duration of classroom events (lecture, clicker questions, group discussion, etc.) over one academic year. Of particular interest was how variable the use of individual as well as group active learning was in the courses comprising a typical introductory biology sequence.

The specific objectives of this study were to:

1. **Document the occurrence of and time spent on different classroom events.**
2. **Compare faculty implementation of individual and group active learning in their classrooms.**
3. **Examine the overall implementation of individual and group active learning across the sequence of courses.**

Methods

Participants and Courses

Participants for this study included eight of the nine instructors of two introductory biology majors' courses at a large southern public university during the 2012-2013 academic year. The two courses, Biodiversity and Cell Biology, seated between 150 and 225 students per lecture section, each of which was taught by a single instructor. The number of lecture sections was two to three per course per semester. These two introductory courses comprised the typical first-year requirement for biology majors at this university. Each course was four credits and also included a one-credit guided inquiry laboratory, which was taught by graduate teaching assistants (GTAs). The laboratory portion of all courses used the same curriculum, which focused on the development of scientific investigation in the context of the course topics, but did not match the lecture content week by week. The laboratory portion of the class was not included in this study, and GTAs did not contribute instructionally to the lecture. One instructor had been involved in development of the labs, and another was involved in coordinating lab delivery, but otherwise, faculty involvement in the labs was minimal. The lecture courses' curriculum was in the process of being reformed to center around the *Vision and Change in Undergraduate Biology Education* final report (AAAS, 2011) concepts and competencies. At the time of the study, the course instructors covered similar agreed-upon topics in their classes, but were free to individually choose the sequence and details of those topics; there was not a unified syllabus in place for each course.

Informed consent to participate in classroom observations was sought and obtained before any data collection began. All participating instructors were observed in the semester(s) in which they taught. All instructor names reported in this study are pseudonyms and the second author was also observed as part of the study.

Data Collection

Classroom teaching observations were used to document different classroom events in the introductory courses. Observations were done using an observation protocol created via an iterative process of classroom observation and event categorization by the first author. Existing observation protocols were also tested, but at the time, no other protocol existed to capture the data needed in this study (e.g., COPUS; Smith, Jones, Gilbert, & Wieman, 2013). Free-hand notes were ultimately found to be the optimal way to capture the classroom events relevant to this study; they were used to log the duration of time spent on each event and, in the case of active learning events, whether it was conducted as an individual or group activity. All classroom observations for this study were done in person, not video recorded. To confirm the reliability of the created observation protocol, an external evaluator was trained in the use of the protocol and observed four classes at the same time as the first author. The duration of classroom events were then compared. Both observers classified all classroom events into the same categories. The duration of the events in all observations matched exactly or within one minute.

There were seven categories of classroom events that were recorded in this study: clicker questions, instructor verbal questions, activities, student questions, lecture, video, and no class (Table 1). The first three events were considered to be

Table 1. Classification of classroom event categories

Classroom Event	Classification
Clicker Questions	Individual/Group Active Learning
Verbal Questions	Individual/Group Active Learning
Activities	Group Active Learning
Student Questions	Student Questions
Lecture	Non-Active
Video	Non-Active
No Class	Non-Active

instructor-led active learning practices and could also be sub-categorized as individual active learning or group active learning. Student questions were placed into a separate category because, although they are a form of active learning, this study was focused solely on instructor use of active learning. Lecture, video, and no class were classified as non-active events. No class is a category that represents the percentage of time when a class begins late or lets out early. Five of the categories were created prior to testing of the protocol, based on previous classroom observations. The activities and video categories emerged as new data collection categories during class observations, and resulted in revision of the observation protocol. Activities all included the use of paper, for example, students being asked to analyze data and turn the work in on paper, or being given pre-created worksheets to guide their discussion. Activity topics included students creating hypotheses, predicting future phenotypes, creating population growth curves, drawing a food web, graphing data, and creating phylogenetic trees. Any clicker questions or verbal questions that were not conceptual in nature (e.g. feedback about a topic or asking about class logistics) were not counted in the active learning total.

In summary, the only classroom events that counted toward active learning in a class were individual or group conceptual clicker questions, individual or group instructor conceptual (verbal) questions, and class activities (which were always implemented using group active learning pedagogies in this study). The observer recorded the length of all classroom events for the entire scheduled class period. At no point was a judgment ever recorded about the quality of the classroom event, e.g., whether certain clicker questions were more effective at fostering conceptual learning or whether particular activities seemed to provoke more student discussion. The data therefore,

reflect only instructor implementation of particular classroom events, and not relative quality of those events.

The first author conducted observations of each consenting instructor once per month over the full semester of instruction. The observation days were chosen randomly and were unannounced. The observation lasted for the entire length of the chosen class, which was either 50 minutes or 75 minutes, depending on the lecture day. One instructor taught in both the fall and spring semesters and therefore was observed nine times over the school year. The other instructors taught during only one semester and were observed between three and five times over that semester. The observations were unannounced and the observer typically sat in the back of the class to be as unobtrusive as possible. All procedures were reviewed and approved by the Institution's Review Board for Human Subjects.

Data Analysis

For each instructor, the percentage of time spent on each classroom event was calculated as a function of the entire class period time (50 or 75 minutes) for each observation. These values were then averaged for each instructor to determine the overall percent of class time spent on each classroom event over the semester. These data were also broken down into individual versus group active learning percent time per instructor. A one-way between subjects analysis of variance (ANOVA) was calculated to determine if instructor differences existed in use of individual and group active learning. Post-hoc comparisons using the Bonferonni correction were performed (SPSS, IBM, 2011). Individual instructor data were averaged by course (Biodiversity and Cell Biology) and for the entire biology sequence. To determine if the differences in

both individual and group active learning (DV) were influenced by the duration of class (50 or 75 minutes) and number of observations performed while controlling for instructor differences, an ad hoc analysis of covariance (ANCOVA) was performed. To determine if there were differences in individual or group active learning implementation between the courses, a comparison of the percentage of class time spent in individual and group active learning in Biodiversity versus Cell Biology classes were conducted using separate independent two-tailed *t*-tests in SPSS (IBM, 2011).

Results

All seven classroom events were observed to occur throughout the academic year overall, however not all events were observed to occur in all instructors' classrooms. On average, lecture was used for 61% of class time, videos made up 2%, class let out early or began late 8% of the time, clickers were used 10% (5% individual and 5% group), verbal questions made up 9% (8% individual versus 1% group), group activities composed 5%, and 4% of class time was spent with students asking questions (Table 2). Clicker questions were the active learning pedagogy used most often, and were implemented using group active learning most often for most instructors. All instructors with the exception of Harold used clicker questions with group discussion. Only Celine used verbal questions with student discussion. Four instructors (Albert, Celine, David, and Erin) used group activities for 8% - 19% of class time, but Erin was the only instructor that spent more time using activities than any other active learning strategy. Average individual active learning use per instructor ranged from 5% - 28% of class time and average group active learning use ranged from 0% - 36% with only one

Table 2. Average proportion of time spent in classroom events by instructor.

Instructor	Clickers	Verbal	Activities	Student Questions	Lecture	Video	No Class
Albert	0.17	0.02	0.08	0.08	0.52	0.00	0.13
Bruce	0.17	0.18	0.00	0.05	0.55	0.03	0.02
Celine	0.07	0.10	0.09	0.09	0.57	0.07	0.01
David	0.07	0.19	0.15	0.02	0.49	0.00	0.08
Erin	0.14	0.07	0.19	0.05	0.55	0.00	0.00
Gail	0.03	0.03	0.00	0.01	0.82	0.05	0.06
Harold	0.06	0.01	0.00	0.00	0.61	0.05	0.27
OVERALL	0.10	0.09	0.05	0.04	0.61	0.03	0.08

instructor not using group discussion (Figure 1). There were significant differences among instructors in the time they spent on individual active learning pedagogies, $F(7,28) = 3.75$, $p < 0.01$, $\eta^2 = 0.48$. Albert ($M = 0.32$, SD = 0.16) spent 6 times more class time using individual active learning than both Gail ($M = 0.05$, $SD = 0.03$ $p < 0.01$) and Harold ($M = 0.06$, $SD = 0.02$ $p < 0.05$). There were also significant differences between instructors in the time they spent on group active learning, $F(7, 28) = 3.55$, $p < 0.01$, $\eta^2 = 0.47$. The only significant pairwise comparisons were comparing Celine to Gail and Harold; Celine ($M = 0.24$, $SD = 0.07$) spent almost a quarter of her class time on group active learning, while Gail spent only 1% ($M = 0.01$, $SD = 0.02$, $p < 0.05$) and Harold spent none ($M = 0.0$, $SD = 0.0$). There were no significant effects of number of observations and length of class on the active learning differences among instructors.

The eight instructors teaching the two introductory biology courses spent 26% of class time on active learning for the overall course sequence, with 16% of class time spent on individual active learning pedagogies, and 10% of class time on group active learning pedagogies (Figure 2). Overall, group active learning made up 38% of all active learning that occurred in the classroom. Biodiversity instructors ($M = 0.23$, $SD = 0.15$) spent significantly more time in individual active learning when compared to Cell Biology instructors ($M = 0.08$, $SD = 0.07$, $t(34) = -3.643$, $p < 0.01$). Biodiversity instructors (M = 0.14, SD = 0.13) also spent significantly more time using group active learning than Cell Biology instructors (M =0.05, SD = 0.1, $t(34) = -2.238$, $p < 0.05$; Figure 2). Despite these differences, the ratio of time spent within group active learning compared to time spent in individual active learning did not differ between courses, with Biodiversity

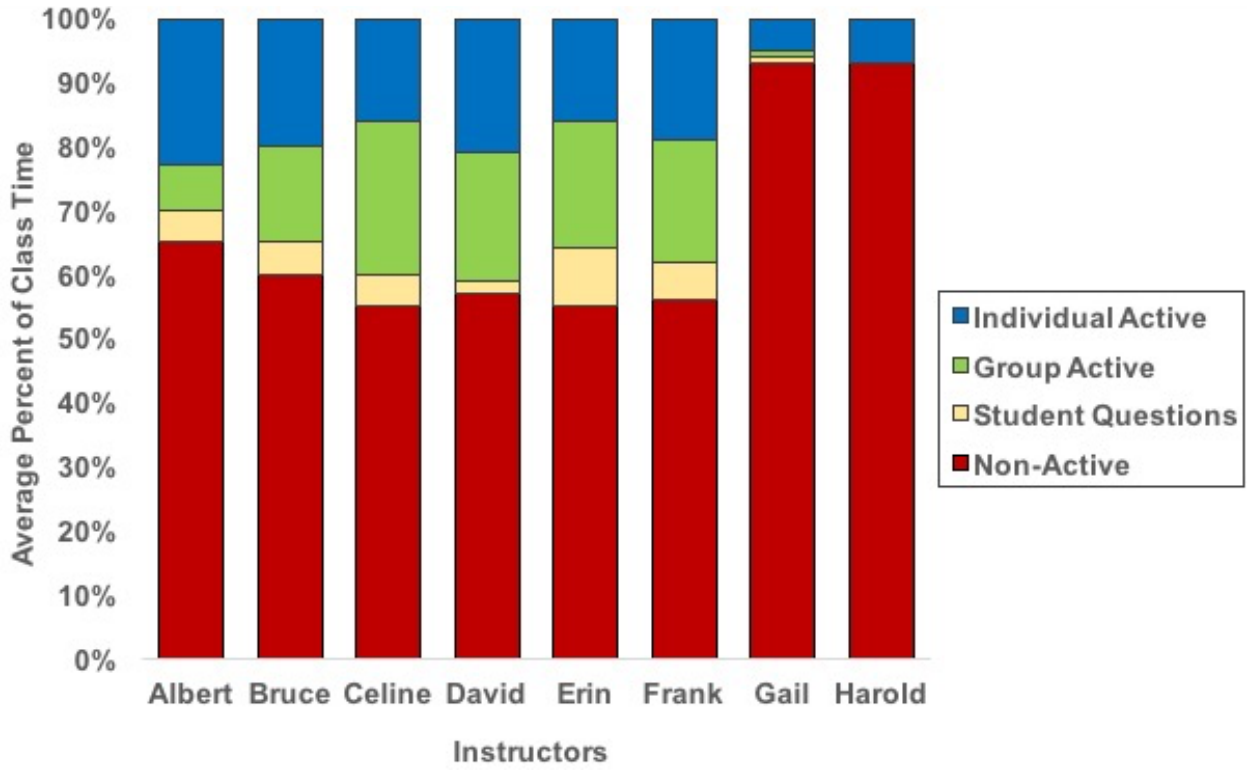


Figure 1. Average event type by instructor

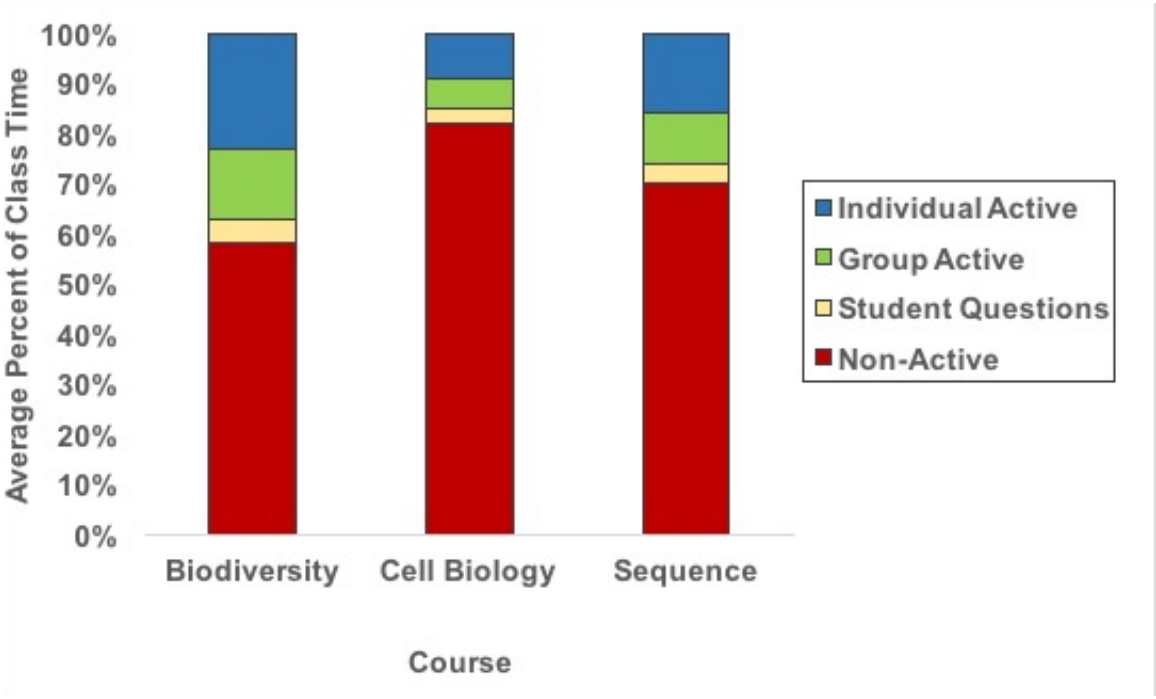


Figure 2. Average Percent Class Time in Classroom Event by Course

instructors spending 38% of total active learning time within group pedagogies and Cell Biology instructors spending 36% of total active learning time within group pedagogies.

Discussion

For the academic year studied, all the instructors of the introductory biology sequence at this university used at least some active learning pedagogies in their classrooms, meaning that their students likely spent at least some time explicitly thinking about course concepts during class time. However, not all instructors implemented active learning using a collaborative or group context, and there were significant differences between the introductory courses in use of active learning. Instructor differences in the amount of group active learning were also reported in a study by Turpen and Finkelstein (2009). Given what research has revealed about learning in small groups, and the utility of self-explanations and student talk (Chi et al., 1994; Smith et al., 2011; Webb, 1989), these differences could potentially result in different student learning outcomes from these courses.

Smith et al. (2014) suggested that observational data be used as a guide for the professional development of faculty. In the case of this course sequence, the instructional faculty began meeting as a community in 2012-2013 to plan for the curriculum changes that were starting in fall 2013. Meetings occurred every two weeks to discuss how to enact changes consistent with the *Vision and Change* report, and professional development was a natural part of these meetings. Professional development during the year of the study was focused on discussing the value of active learning for student learning, not on specific pedagogies. The baseline data reported in this study were used as feedback to the faculty allowing them to see where they stood

in comparison to their own perceptions, others teaching the course, and the sequence as a whole. Talking with each other and sharing personal reflections about teaching was critical to this process, and helped many faculty who were more tentative about making changes than others within the group. The data from this study informed future professional development by revealing the need to integrate more group active learning in lecture courses to increase the probability of learning gains for students. The following year, professional development concentrated on how group active learning could be implemented in introductory lecture courses.

Bringing the results of studies such as these to the attention of program faculty may be useful, because it is likely that they may not realize how little group active learning is being used. Ebert-May et al. (2011) found disconnects between faculty's perception of their teaching and observed practices. Instructors may also not realize how important discussion is to student learning. Once understood, faculty can convert their individual active learning strategies to group discussion strategies, which may increase the potential for student learning from these introductory courses. Our experience from this study indicates that this needs to happen within the context of a community of faculty sharing and discussing teaching experiences.

Overall, clickers were the most commonly used individual active learning and group active learning strategy observed in the classrooms. They are, perhaps, the easiest entry into active learning for most instructors as the grading of responses is automatic. Clickers could also be the easiest way to integrate *group* active learning into these courses by providing professional development to faculty members about the types of clicker questions and types of instructional prompts they can use to foster

student discussion. Similarly, verbal questions may also be a way to increase active learning in college classrooms given their frequent use by the instructors, although our study found that they were rarely used to foster group discussion. Pairing verbal and clicker questions with discussion not only increases fluency within a given topic (Herman, Munk, & Engel, 2004; Shaw, McGaugh, & Rose, 1990), but also improves student understanding even when none of the students know the correct answer (Smith et al., 2009). However, instructors should also discuss within their faculty communities times when non-active learning, individual active learning, and group active learning are best implemented. Each pedagogy has a role to play within the classroom, and instructors can use the principles of backward design (Wiggins & McTighe, 1998) to make decisions about how to maximize student understanding of each class learning objective.

Only the instructors of Biodiversity used pre-planned activities in the classroom. The activities were all group active learning and included experimental design, predicting phenotypes, constructing phylogenies, and working with scenarios. These activities take more forethought to integrate into class time, more class time to implement, and a longer time to evaluate for feedback so it is not surprising that they were used less compared to clickers. Their benefit, however, is that they inherently increase group active learning in the classes. The use of cooperative learning in the classroom has been shown to contribute to students putting in greater effort to achieve success, as well as improve the quality of peer relationships and promote psychological well being (Johnson & Johnson, 1989). As observations of these courses continue into the future, instructors who currently only use clickers and verbal questions will be

tracked to see if they will adopt the use of activities as well. This might suggest that activities are integrated only after initial testing of lower-risk active learning strategies such as clicker use.

Instructors of the Biodiversity course were more likely to implement different amounts and types of active learning, including group active learning, in their classes than instructors of the Cell Biology course. It is not known if these course differences are due solely to the nature of the individual instructors or to a difference in the content or culture of the two courses. The instructors for the Biodiversity course are from an ecology and evolution department, while those who teach the Cell Biology course are from a cell and molecular biology department. Different sub-disciplines of biology may form different communities of practice that impact the way they view instruction and student learning (Lemke, 2001). This difference was not anticipated and it may be important for continued research to investigate sub-disciplinary perspectives on the use of active learning to identify possible reasons for these differences. Topics in cell biology courses are often more abstract and less visually tangible than topics in biodiversity perhaps resulting in instructors having more difficulty creating activities or questions that students can easily discuss. Therefore, the same active learning techniques, such as self-explanation, might not be appropriate or successful for both courses and could be a reason instructors teaching these two courses differ.

One limitation of this study is the low number of observations per instructor. Only 8 – 20% of each instructor's classes were sampled over the semester. Although there was no evidence based on our statistical analyses that this sampling impacted the results, small sample sizes could paint a skewed picture of what is actually occurring in

the classroom. However, this sampling regime was able to capture a diversity of classroom active learning pedagogies and in discussing the data, instructors generally felt that the results represented their classroom practices. Another limitation is that quality of active learning was not considered for this study. This could lead to a misinterpretation of active learning effectiveness because time of active learning use may not necessarily correlate with quality of active learning event. However, we did find that instructors who spent more of their class time in active learning were also more likely to use active learning practices such as activities or group clicker questions, which are known to increase the probability of student self-explanation (Chi et al., 1994; Webb, 1989). Therefore, an area for further research may be the relationship between quantity of active learning use and the quality of those practices.

The findings from this study provided an overview of pedagogical practices within these introductory courses that was invaluable to our curriculum reform efforts. The goal is for these data to help instructors to increase the use of group active learning through discussions about specific practices that may be used in these particular classes. It is important for all institutions to conduct assessments of current and reformed practices to collectively meet the national goals articulated in the *Vision and Change* report (AAAS, 2011). This collective research will provide a better understanding on a national level of how to foster change in STEM classrooms, and how to support faculty in their planning and implementation of group active learning. In the meantime, students will likely experience a patchwork of non-active, individual active learning, and group active learning practices in their introductory courses, which may lead to a patchwork of conceptual understanding of biological concepts in these critical gateway courses.

CHAPTER II

A VISION AND CHANGE REFORM OF INTRODUCTORY BIOLOGY IMPROVES FACULTY PERCEPTIONS AND USE OF ACTIVE LEARNING

Abstract

Helping to increase faculty use of active learning pedagogies in college classrooms has been a persistent challenge in biology education. The University of Tennessee, Division of Biology, implemented curriculum changes to their majors' two-course introductory sequence as outlined by the *Vision and Change in Undergraduate Biology Education* final report (AAAS, 2011). Goals of the curriculum reform included integrating core biological concepts and competencies into the courses using active learning pedagogical approaches. The purpose of this study was to observe the instructional practices used by faculty (N = 10) throughout the process of reform to determine if the use of active learning strategies increased. Instructors also participated in interviews to track their perceptions of instruction as the reform progressed. Instructors increased their average use of active learning by 12% throughout the three-year study. Collaboration with other instructors, feedback with reflection, and time were found to be necessary elements for instructional change to be achieved. Evidence is provided that small shifts in instruction over time can accumulate into real change in the classroom.

Introduction

It has been posited that the traditional lecture format, as opposed to more active learning (AL) formats, of undergraduate science courses presents many challenges for student learning overall. While lecture is an efficient way to disseminate a large amount of information to a large number of students, this teacher-centered strategy promotes passive student learning (Bransford, Brown, & Cocking, 2000) while stifling student motivation and enthusiasm (Weimer, 2002). As evidence of this, many students who leave science majors say it was because they felt uninspired and perceived the

instructors as not caring (Seymour & Hewitt, 1997). Roughly 60% of STEM undergraduate students, including 58% of life sciences students, switch their majors or dropout of university; this percent for life sciences students increases to 80% for women and students from minority groups (PCAST, 2012). Active learning (AL) is a student-centered pedagogical approach that engages student thinking through the use of class activities that require students to reflect upon their ideas and their application (Collins & O'Brien, 2003). Introductory biology courses have found AL to improve student performance, including underrepresented students (Armbruster, Patel, Johnson, & Weiss, 2009; Freeman, et al*.,* 2007; Michael, 2006), and increasing the use of student-centered practices is a main recommendation of the *Vision and Change in Undergraduate Biology Education* final report (AAAS, 2011). Freeman et al. (2014) recently suggested that it is time to stop advocating for the use of AL and instead begin implementing it.

There are various instructional methods that can be used to incorporate AL in the classroom, with common practices including clicker questions, verbal questions, and activities. Many instructors use student response systems or "clickers" that allow large classrooms of students to select their answer for a projected problem or question using a handheld personal response device. Instructors also ask questions out loud, called verbal questions, to produce responses either through student volunteers or calling on students to respond. Writing is another form of eliciting responses from students. Instructors can use classroom activities, such as figure interpretations, model drawings, phylogenetic trees, and ask students to respond in writing on notebook paper, notecards, or pre-printed worksheets.

Active learning (AL) techniques can be used to engage students individually or in a collaborative context. While both have been shown to be successful strategies for increasing student learning, combining peer discussion with instructor explanation has been shown to maximize student learning gains resulting from the use of AL use in biology classes (Smith et al., 2009; Smith et al., 2011). As such, providing opportunities for student discussion is seen to be critical for a student-centered course (AAAS, 2011). Couch et al. (2015) suggests that peer interaction is one component of best practices for scientific teaching, specifically allowing students to refine their knowledge through activities such as worksheets that require small group discussion. Eddy, Converse, and Wenderoth (2015) found that students explaining their answers to peers, hearing other students describe their logic, and participating in activities that require group work are all elements of best practices in AL. Research in physics has suggested that although classroom discussion is important for student understanding (Turpen & Finkelstein, 2010), how discussion is employed in the classroom varies by instructor, with some professors promoting much higher levels of student discussion than others (Turpen & Finkelstein, 2009).

Helping to increase faculty use of AL pedagogies in college classrooms has been a persistent challenge in biology education. It has been found that a one-time workshop is not enough to invoke meaningful change in faculty use of evidence-based practices (Sunal et al., 2001). Faculty who attended a series of professional development workshops to promote AL have been shown to self-report a higher use of AL than assessed in their actual classroom teaching (Ebert-May et al., 2011). Gormally, Evans, and Brickman (2014) highlight the importance of supportive instructional feedback as

faculty attempt to implement AL strategies in their classrooms. Even when faculty indicate that they are familiar with the literature describing successful pedagogical practices, they still report that it has not led to significant changes in their classroom. Faculty often report that time for preparation is the largest barrier for implementing AL practices (Dancy & Henderson, 2010). It has been suggested that faculty are more likely to put effort into pedagogical reform if they understand the goal, are committed to it, and believe in its success (Kluger & DeNisi, 1996). Given this, a process that incorporates the collaborative development of instructional goals among the faculty implementing the change and providing instructional feedback during implementation may be beneficial to enacting pedagogical change.

The study took place over a period of three consecutive years as the instructors of an introductory biology sequence at a southeastern public research university implemented curriculum changes guided by the *Vision and Change* final report (AAAS, 2011). This study built on previous research examining AL use in introductory biology at the same institution (Auerbach & Schussler, in press). One of the goals of the curriculum reform was to use evidence-based instructional practices, such as AL, as a means for delivery of the new curriculum. The purpose of this study was to track faculty perceptions and use of AL instruction in the introductory biology majors' courses throughout the three-year curriculum reform process using classroom observations and interviews. These combined data (quantitative and qualitative) were used to answer the following research questions:

1. **How are active learning practices changing within the program throughout curriculum reform?**

2. **How are active learning beliefs changing within the program throughout curriculum reform?**

Methods

Participants

The participants for this study included instructors (N = 15) who taught one of the two introductory majors' biology courses (Biodiversity and Cell Biology) between fall 2012 and spring 2015. Because the intent of the study was to track instructional change over time, instructors were not included if they did not teach at least twice over the three years of the curriculum reform period. This reduced the number of instructors to ten; five instructors taught the Cell Biology course and five taught the Biodiversity course (Table 3). All ten instructors taught at least once a year for two years with four instructors teaching all three years. All of the instructors held a PhD, half of the instructors were tenure-track faculty, and the other half were full-time permanent Lecturers. Most of the instructors (N = 8) had taught the introductory course for less than five years, two had 5-10 years of experience with the course, and one had more than 10 years of experience in the course; only one instructor who participated in this study was new to either course at the beginning of the study (Table 3). Informed consent was received from all participating instructors for both observations and interviews prior to collecting data. All observation procedures were reviewed and approved by the University of Tennessee Review Board for Human Subjects.

Courses and Process of Change

The Biodiversity and Cell Biology courses were offered in both fall and spring semesters. There were eight total lecture sections observed each year with four

Table 3. Summary of instructor demographics.

Summary of instructor pseudonyms, course taught, experience with course, and the years they were observed and interviewed for this study.

Pseudonym	Course	Teaching experience in course	Years Observed/Interviewed
Albert	Biodiversity	< 5 years	Year 1 Year 3
Bruce	Biodiversity	10+ years	Year 1 Year 2 Year 3
Celine	Biodiversity	< 5 years	Year 1 Year 2 Year 3
David	Biodiversity	< 5 years	Year 1 Year 2
Erin	Biodiversity	< 5 years	Year 1 Year 2 Year 3
Frank	Cell Biology	5-10 years	Year 1 Year 3
Gail	Cell Biology	< 5 years	Year 1 Year 2
Harold	Cell Biology	< 5 years	Year 1 Year 2 Year 3
Ingrid	Cell Biology	< 5 years	Year 2 Year 3
Juanita	Cell Biology	5-10 years	Year 2 Year 3

representing Cell Biology and four representing Biodiversity. Each lecture section contained 150-225 students. There were honors lecture sections for each course that were not included in this study, and the summer offerings of each course were also not included in the study. The Biodiversity course is the first in the sequence and covers topics such as ecology, evolution, and the diversity of life. The Cell Biology course covers topics such as macromolecules, cell structure, metabolism, and DNA structure and function.

The curriculum changes undertaken were guided by the *Vision and Change in Undergraduate Biology Education* (AAAS, 2011) final report. Select faculty members across the three biology departments served by the introductory courses formed a task force in 2010 to consider ways to improve the courses. After the publication of the *Vision and Change* final report (AAAS, 2011), the faculty decided to adopt the concepts and competencies from the report as the new unified learning objectives for the two courses. A decision was also made to remove the labs appended to each course and instead add one weekly 50-minute small group discussion led by teaching assistants (TA) focused on primary literature and biological literacy in an AL context. This resulted in a proposal to switch from a traditional two course 8-credit introductory sequence to a three course 8-credit sequence with two lecture/discussion courses (3 credits each) and an independent 2-credit hour lab and discussion starting in fall 2014. This research project focused only on the lecture portions of the four credit courses that existed through spring 2014, and then the first delivery of the two newly-created 3-credit lecture courses in the 2014-2015 academic year. All lecture content (core concepts) in the courses remained the same, but the structural changes to the new courses resulted in

50 minutes less lecture time per week than in the original courses (150 minutes per week in the old courses versus 100 minutes per week in the new courses).

To ease the transition to the new course structure, the curriculum changes were phased in over three years of reform. During the first year (2012 – 2013), four faculty members piloted the *Vision and Change* concepts and competencies within their lecture courses. During the second year (2013 – 2014), all faculty used the *Vision and Change* concepts and competencies in their courses, but the course structure remained as a 4-credit lecture-lab format. During the third year (2014 – 2015), the new course structure with 3-credit lecture plus discussions and one two-credit lab with discussion were implemented.

Throughout the curriculum reform process, faculty met as a group to discuss the reform, participate in professional development sessions, and share resources. Instructors worked together and jointly developed the implementation guidelines for the lecture courses. Faculty also discussed course data (such as AL use) and pedagogy throughout the process. On average, these meetings occurred once a month and lasted at least one hour each. At the end of each semester of observations, the instructors had access to their own AL data as well as the compiled data for the program. All data were shared anonymously. Some of the instructors also participated in communities of faculty, postdoctoral associates, graduate students and undergraduates who worked together to design the discussion curricula (Dalrymple, Auerbach, & Schussler, in prep), however, the meetings with instructors about changes to the lecture courses were separate from the meetings about the new discussion curricula.

Data Collection

Classroom observations were used to sample instructional practices within the introductory courses over the three years of the reform process. Observations were undertaken using an observation protocol created via an iterative process of classroom observation and classroom event categorization by the first author in early fall 2012. Existing observation protocols were also tested, but at the time data collection began, no other protocol existed to capture the data considered central to this study (COPUS; Smith, Jones, Gilbert, & Wieman, 2013, had not yet been published). After the publication of COPUS, the authors discussed switching to that instrument, but since the switch would not necessarily have resulted in improved data for the study, the decision was made to keep the data collection process the same. Free-hand notes were found to be the optimal way to capture the classroom events relevant to this study; they were used to log the duration of time spent on each event and, in the case of AL events, whether they were conducted as an individual or group activity. All classroom observations for this study were done in person, not video recorded. To confirm the reliability of the created observation protocol, three external evaluators were trained in the use of the protocol and observed at least two of the same classes as the first author each semester. Observers classified all classroom events into the same categories, and the duration of the events in all observations matched exactly or within one minute of each other.

The first author conducted observations of each consenting instructor once per month over the full semester of instruction for three academic years. The observation lasted for the entire length of the chosen class, which was either 50 minutes or 75

minutes, depending on the lecture day (50 minutes for courses taught on Mondays, Wednesdays, and Fridays vs 75 minutes for courses meeting on Tuesdays and Thursdays). The observations were unannounced and the observer typically sat in the back of the class to be as unobtrusive as possible. In year 3, observations were also conducted for the newly implemented discussion sections. There were a total of three graduate teaching assistants (TAs) assigned to each lecture course, and each of those TAs taught three discussion sections with up to 25 students each. Each TA was observed twice during each semester they taught using the same protocol as the lecture observations.

There were seven categories of classroom events that were recorded throughout the study: clicker questions, instructor verbal questions, activities, student questions, lecture, video, and no class. These categories were classified as either AL or non-AL. Classroom events were classified as AL if the instructor asked the students to engage with the material in some manner (i.e. answer a clicker question, discuss a concept, draw a phylogenetic tree). Active learning was also broken into sub-categories of individual (students work alone) and group (students work together) AL. The classroom events of clickers, verbal questions, and activities were considered to be instructor-led AL practices and could also be sub-categorized as individual AL or group AL. Any clicker or verbal questions that were not conceptual in nature (e.g. feedback about a topic or asking about class logistics) were not counted in the AL total. Student questions were placed into a separate category and not considered AL because in this case the instructors were not explicitly asking students to be engaged in their learning and this

study was focused solely on instructor-directed AL. Lecture, video, and no class (class beginning late or dismissed early) were classified as non-AL events.

Five of the AL categories were created prior to testing of the protocol, based on previous classroom observations. The activities and video categories emerged as new data collection categories during class observations, and resulted in revision of the observation protocol. Activities included the use of paper, such as students being asked to analyze data and turn the work in on paper, or the use of worksheets to guide discussion and responses. Activity topics included students creating hypotheses, predicting future phenotypes, creating population growth curves, drawing a food web, graphing data, and creating phylogenetic trees. In summary, the only classroom events that counted toward AL in a class were individual or group conceptual clicker questions, individual or group instructor conceptual (verbal) questions, and individual or group class activities.

The observer recorded the length and frequency of all classroom events for the entire scheduled class period. At no point was a judgment ever recorded about the quality of the classroom event, for example, whether certain clicker questions were more effective at fostering conceptual learning or whether particular activities seemed to provoke more student discussion. The data therefore reflect only instructor implementation of particular classroom events, and not relative quality of those events.

Each instructor was interviewed at the end of the semester they taught. The same questions (Appendix) were asked at each interview in order to identify changes in perception and use of active learning over time. The interview questions were designed based on a priori categories (course planning, AL implementation, AL definition, AL

types and teaching style) to track changes over time in the way faculty thought about, planned for, and implemented AL within their classrooms. For example, faculty were asked what active learning means to them and what types of active learning they use in their classroom. Follow-up questions were often used to clarify or extend participant responses. The interviews were either held in the participant's or first author's office depending on the participant's preference. The first author performed all interviews. The interviews were audio-recorded with permission; beginning only after participants received all information about the study and signed an informed consent form. The interviews ranged from 18 – 36 minutes each. Instructors were interviewed twice if they taught two of the three years and three times if they taught every year, resulting in a total of 24 interviews with the 10 participants over the course of the project.

Data Analysis

While fifteen instructors were observed, only the ten core instructors were included in data analyses. During the first year of observations, the number of observations per instructor varied from three to five. In order for each instructor to have equal representation and for each year to be compared over time, it was determined that each instructor would have three observations in the fall (one each in September, October, and November) and three in the spring (one each in February, March, and April). Thus, for the first year of data collection, three random observations were selected within each month for each instructor each semester to standardize the data analysis (Table 3). Eight of the ten instructors taught each year, thus, a total of 24 observations were analyzed each year for a total of 72 observations over the course of reform.

Observational data were used to calculate a proportion of class time each instructor spent in each classroom event per observation and was recorded into a spreadsheet for analysis. The percentage of time spent on each classroom event was calculated as a function of the entire class period time (50 or 75 minutes). These values (N = 24 per year) were then averaged across instructors to determine the overall percent of class time spent on each classroom event per year. AL classroom events were clickers, verbal questions, and activities. The values that represented the three AL events were added together for each instructor to produce the total proportion of class time spent in active learning overall (ALO). These values were also broken into individual and group AL events based on the implementation (i.e. students answer a clicker individually versus after discussing the question with classmates) to produce a measure of group active learning overall (GALO) for each instructor. These values were then averaged across instructors to determine the overall percent of class time spent on each classroom event (clickers, verbal questions, and activities), ALO, and GALO over each of the three academic years. The data were assessed for coding errors and missing data, then standardized scores were created to check for outliers. A Shapiro-Wilk test was used to assess the data for non-normality as the sample size was less than 300. Appropriate skewness and kurtosis were also confirmed (Westfall & Henning, 2013).

A one-way repeated measures analysis of variance (ANOVA) using the observation data from lecture only was performed with the total amount of active learning (ALO) as the dependent variable and year as the independent variable to determine if instructor use of AL was changing across the reform. Separate one-way

repeated measures ANOVAs were performed to determine if changes occurred across time in use of group active learning (GALO) or classroom event (i.e. clickers, verbal questions, and activities). For analyses that produced significant results, post-hoc comparisons using the Bonferroni correction were performed to determine which years were different (SPSS, IBM, 2011).

To compare the 150 minutes of lecture per week in the first two years to the 150 minutes of lecture plus discussion per week in the third year, the discussion section observation data were added to the lecture observation data in 2014-2015 to come to a final value that represented the total AL students would experience in the 150 minute course per week. Observations of TAs were transformed into proportions of class time spent in different classroom events. Active learning overall (ALO) and GALO were calculated for each TA observation. All TA values associated with the same section of a course were averaged together. These data were then used to represent 1/3 of the observational averages per class, with the instructor observations representing the other 2/3.

The addition of the discussion data to the lecture data transformed the dependent variable into one that violated sphericity because the use of AL in the discussion sections was significantly higher than the values seen in lecture. Sphericity is an important consideration when using repeated measures ANOVA as it can affect the accuracy of the ANOVA *F* statistic by reducing its power (i.e., type 1 error). To correct for this, a one-way repeated measures ANOVA was calculated using the more conservative Greenhouse-Geisser correction (Stevens, 1996). For analyses that

produced significant results, post-hoc comparisons using the Bonferroni correction were performed to determine which years were different (SPSS, IBM, 2011).

Although fifteen instructors were interviewed over the course of the project, only ten were included in this study, and only a combination of eight out of those ten taught each year of the project. Since the intention was to analyze interview responses for potential changes on a yearly basis, the interview analyses were carried out on only the eight instructors who taught each year. This resulted in the initial pool of 29 interviews being reduced to 24 for the purposes of data analysis. The 24 interviews were transcribed fully into a word processing program for analysis. The interviews were then analyzed in two successive cycles using two independent researchers, one being the primary analyst and the other independently confirming the themes identified by the first.

In the first round of analysis, interviews were analyzed using the a priori categories that were developed and reflected in the interview questions. These categories were used as a guide to examine the responses and identify themes within each category in order to search for change across time (Kvale & Brinkmann, 2009). To do this analysis, all responses relating to each category were highlighted. Responses for each category were read and notes were made about potential themes. The responses were then re-read and initial sorting into themes was made to determine if they worked for the data. Then themes were finalized when it was apparent that they represented the majority of the responses (Saldaña, 2013). These themes were then used to create a summary of each category by tallying the number of instructors whose responses could be placed in each theme; for example, under the category of course

planning, instructors reported two themes of behavior: using previous course materials as a guide for course planning or developing learning objectives for the course to serve as the guiding structure.

After this first round by the primary analyst, a co-researcher in a subsequent round also assessed the interviews. The co-researcher identified participant responses using the a priori categories and themes created by the primary researcher. The co-researcher also made notes of possible a posteriori categories and themes that were not planned for or expected but emerged from the data nonetheless. Once the co-researcher sorted the participant responses under each category by the themes, the two researchers met to compare results and reach consensus. At this time, some themes were re-titled or compacted, but there were no significant disagreements about the presence of the created themes. However, there were two categories, in addition to the a priori categories, that the researchers felt warranted further exploration (assessment and professional development). Each researcher independently searched the interviews for participant responses that related to the two new categories. Then they each assigned the responses to a theme within each category. Finally, they met again to compare results and reach consensus. Assessment was added as a category, but it was decided that there was not enough about professional development to warrant a category addition. Summaries for each category and theme were then compared across the years using comparative analysis (Saldaña, 2013). For example, the number of faculty using previous course materials as a guide for course planning decreased over the three years.

Results

Observations

Based on percent class time alone, instructors increased their average use of AL, group AL (Figure 3), as well as all three AL strategies throughout the three-year study (See Table 4). During the first year, instructors spent an average of 26% of their class time in AL with 13% spent in clickers, 8% in verbal questions, and 6% in activities. Group AL was used for 13% of the overall AL (Table 4). During the second year, overall AL use was 27% of class time on average. Clickers made up 11%, verbal questions 6%, and activities 9% of class time on average. Group AL made up 15% of the overall AL. During the third year, when looking at the lecture only data, 39% of class time was spent in AL events, on average. Group AL accounted for 21% of overall AL. Clickers were used 17% of the time, verbal questions made up 10%, and activities were used 12% of the time. AL use did not decrease from year 2 to year 3, despite a loss of 1/3 instructional time. When looking at the lecture plus discussion data for year three, overall AL made up 49% of class time with 33% representing group AL. Clickers were used for 13% of class time, verbal questions made up 20% and activities 26%.

Over the three years of reform, there was an overall significant effect of ALO use when comparing lecture observations only, $F(2, 46) = 3.95$, $p = 0.026$, partial $\eta^2 = 0.147$, power = 0.682. The results of a bonferroni posthoc test showed no significant results for paired year comparisons: year 1 ($M = 0.26$, $SD = 0.18$), year 2 ($M = 0.27$, $SD = 0.17$), and year 3 (0.39, SD = 0.14). However, comparing year 2 and year 3 resulted in marginal significance, $p = 0.07$. Although the percentage of class time devoted to AL classroom events increased over time, there were no significant differences when

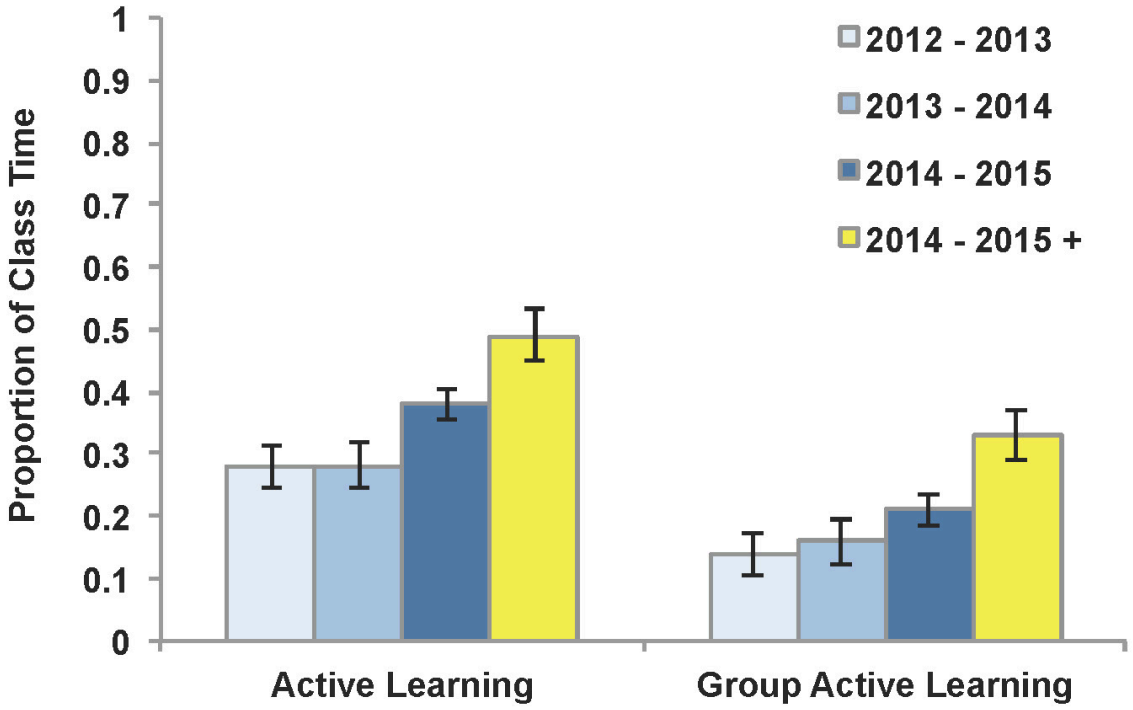


Figure 3. Proportion of class time in active learning and group active learning across three years of reform with 2014 - 2015 shown with and without discussions included.

Table 4. Average proportion of class time in active learning across reform.

Average proportion of class time spent in different types of active learning in all introductory courses in each of the three years of reform. Active Learning (ALO) was the total of all types of active learning, while Group Active Learning (GALO) was only when students were allowed to discuss clicker questions, verbal questions, or work together on activities. Year 3 is broken down into active learning in lecture classes only, versus active learning in both lecture and the new discussion class added that year.

	YEAR 1	YEAR 2	YEAR 3	YEAR 3 w/discussions
Active Learning (ALO)	0.26	0.27	0.39	0.49
Group Active Learning (GALO)	0.13	0.15	0.21	0.33
Clicker Questions	0.13	0.11	0.17	0.13
Verbal Questions	0.08	0.06	0.1	0.20
Activities	0.06	0.09	0.12	0.26

comparing the use of GALO or any of the types of AL events (i.e. clickers, verbal questions, and activities) in lecture across the three years.

When discussion section data were included in the analyses, there were significant increases in use of ALO in year 3. A one-way repeated measures ANOVA was significant for AL, $F(2, 46) = 10.651$. $p < 0.001$, partial $\eta^2 = 0.317$, power = 0.985. The results of a bonferroni posthoc test showed that year 3 AL use ($M = 0.49$, $SD = 0.2$) was higher when compared to both year 2 ($M = 0.27$, $SD = 0.17$) and year 1 ($M = 0.26$, $SD = 0.18$). GALO violated sphericity so a more stringent F test was used (Greenhouse- Geisser), $F(1.549, 35.633) = 8.647$. $p = 0.002$, partial $\eta^2 = 0.273$, power = 0.915. A bonferroni posthoc test revealed year 3 had higher group AL rates ($M = 0.33$, $SD = 0.23$) than both year 2 ($M = 0.15$, $SD = 0.19$) and year 1 ($M = 0.13$, $SD = 0.16$). Clickers did not change significantly across the three years of reform even when discussion was included. The use of both verbal questions and activities also violated sphericity so the more stringent Greehouse-Geisser F test was used. The use of verbal questions increased across the three years when including discussion data, $F(1.191, 27.382) = 4.099$, $p = 0.046$, partial $\eta^2 = 0.151$, power = 0.539, although there were no significant pairwise comparisons among years: year 1 ($M = 0.08$, $SD = 0.08$), year 2 ($M = 0.06$, $SD = 0.05$), and year 3 ($M = 0.20$, $SD = 0.28$). Comparing years 2 ($M = 0.06$, $SD = 0.05$) and 3 ($M = 0.20$, $SD = 0.28$) revealed marginal significance, $p = 0.089$. The use of activities also increased over time, $F(1.275, 29.323) = 7.571$, $p = 0.006$, partial $\eta^2 = 0.248$, power = 0.822. A bonferroni posthoc test revealed significant differences, with year 3 ($M = 0.26$, $SD = 0.28$) using significantly more activities when compared to both year 1 ($M = 0.06$, $SD = 0.12$) and year 2 ($M = 0.09$, $SD = 0.14$).

Interviews

For the a priori categories probed by the interview questions (Course Planning, AL Implementation, AL Definition, AL Type and Teaching Style) the identified themes were tallied over time to identify changes in instructor perceptions and use of AL (Table 5). Themes within each category are italicized in the descriptions below.

Course Planning. Within this category, two themes were identified: '*instructor used previous course materials as a guide'* and '*instructor used learning objectives as a guide'*. For instance, some instructors reported using old powerpoint presentations or the textbook as a starting point for their course planning. Other instructors talked about learning goals for the course and how these were used to shape the course planning. In the first year of interviews, over half of the instructors reported '*instructor used course materials as a guide'* to develop the focus of their teaching (N = 5). Only three were '*instructor used learning objectives as a guide'*.

Instructor statements that identified each approach are as follows:

> David said, *"I start with the book and select topics, objectives, and how to teach them."*[*instructor used course materials as a guide*]
>
> Celine described her method as, *"I read sections of the book I have assigned to the students. From there I pick out the material that I think is most important for students to learn."* [*instructor used course materials as a guide*]
>
> Bruce used learning objectives in course planning. He stated,*" I start with specific learning objectives. What do I want them to do? How can I achieve that?"* [*instructor used learning objectives as a guide*]

Table 5. Summary of interview categories and themes.

Frequency of themes for each interview category across three years of reform for instructors that taught each year (N = 8). Categories and themes were identified by qualitative analysis of instructor interviews. All categories were created a priori with the exception of Assessment being created posteriori. Eight instructors were interviewed each year that theme.

Category	Theme	YEAR 1	YEAR 2	YEAR 3
Course Planning				
	instructor used available course materials as guide	5	5	4
	instructor created learning objectives	3	3	4
AL Implementation				
	AL planned after lecture is set	7	7	5
	lecture planned around AL	1	0	2
	AL planned simultaneously with lecture content	0	1	1
AL Definition				
	students engaged, involved, thinking about content	7	8	7
	students interacting with each other	0	4	5
	students as knowledge constructors	1	1	3
AL Type				
	instructor reported using group AL	6	8	8
	instructor did not report using group AL	2	0	0
	instructor reported using clickers	8	8	7
	instructor reported using verbal questions	6	7	7
	instructor reported using activities	5	6	8
Teaching Style				
	yes, changed	4	7	7
	do not know	1	0	0
	no, not changed	3	1	1
Mention Assessment		1	3	6

The next year, similar numbers of instructors fell into the two themes. Five instructors reported '*instructor used course materials as a guide',* while three reported '*instructor used learning objectives as a guide'.* This may account for the similarities between AL use in years 1 and 2. For example:

> Gail described her course planning, "*So when I design a class, I typically go by the topics I want to cover which sometimes match the book, most of the time matches the book, then make the powerpoint."* [*instructor used course materials as a guide*]
>
> Harold reported a similar approach, *"I essentially go through the book or other sources and pick out what I think are the important things and leave out the stuff that I don't think is important, the stuff that clutters, kind of streamline it a little bit."* [*instructor used course materials as a guide*]
>
> Ingrid talked about a different approach, *"Usually I start with what do I want the students to know at the end of the class. Basically, what are my learning objectives and then I try to design."* [*instructor used learning objectives as a guide*]

By the third year instructor responses began to change. Half of the instructors (4) reported '*instructor used course materials as a guide'* while the other half (4) reported '*instructor used learning objectives as a guide'*. This perhaps had some influence on the increase in AL use in year 3. For example:

Juanita continued to use course materials as a planning guide, "*I go through the chapter and select what to keep.*" [*instructor used course materials as a guide*]

Erin revealed a shift in her planning, *"What I've done with this semester is I explicitly backward designed my objectives for each class… I create learning objectives, explicit ones, that were more not just like to understand or to know but to write them the way I want them to know them on an exam."* [*instructor used learning objectives as a guide*]

Albert also stated, *"One thing I did this year that I learned was to put in learning objectives and that helped me a lot, helped me to teach better and be better organized."* [*instructor used learning objectives as a guide*]

Active Learning Implementation. There were three themes identified for when the planning for AL implementation occurred: '*planned after lecture is set', 'planned simultaneously with lecture content', 'planned lecture around AL'.* Most of the instructors (7) reported that when implementing AL, it is '*planned after lecture is set'.* Only one instructor reported they '*planned lecture around AL'.* For example:

Harold said,*"[*Active learning implementation*] is all retrospective. I go back now and try to figure out where to insert a clicker question."* [*planned after lecture is set*]

Erin described how she implements active learning, *"After I get the lecture set, I try to find places to insert clicker questions where I can test them on information from the last lecture or something I just covered."* [*planned after lecture is set*]

Celine stated, "*For each topic or concept I cover, I think about what questions and activities could help students learn the material better than lecturing."* [*planned lecture around AL*]

The next year, the trend was similar as in year 1 and most of the instructors (7) reported when implementing AL, it is '*planned after lecture is set'.* One instructor reported AL was '*planned simultaneously with lecture content'.* For example:

Gail said, *"I design activities on the backend instead of the frontend. I also design questions for examples on the backend instead of the frontend."* [*planned after lecture is set*]

Juanita said, "*So I give like a quick summary, especially for these connecting or continuing that lecture and then I'll go through the lecture, you know the powerpoints, so I use the Pearson powerpoints from the book but I cut down a lot of things and take the pictures from them and add my own but it's flow is mainly from the DVD, the CD they gave us. Then I add clickers.*" [*planned after lecture is set*]

Celine expressed similar ideas*, "Any time I'm covering a concept that is difficult and I want students to get a chance to either think about it or talk about it on their own or whenever there is a natural point for them to do something and me to get feedback on whether or not they understand or whether they have questions, I think that, that should be driving when I do my activities more than just timing but*

yeah so I think both of those things honestly play a role in when I decide to integrate it." [*planned simultaneously with lecture content*]

The third year, two instructors reported that they '*planned lecture around AL'* and one reported AL was '*planned simultaneously with lecture content'.* The remaining instructors (5) reported that AL is '*planned after lecture is set'.* For example:

Ingrid designed around her active learning, *"The one thing I think I centered more about the activities I had, I think a set of activities that I like that worked well in the past. And so I basically designed around this."* [*planned lecture around AL*]

Celine continued to plan AL integration simultaneously with lecture content, *"In addition to just trying to make sure I'm not talking for long stretches of time, I think it's nice to have breaks from that but also any time I'm covering a concept that is difficult and I want students to get a chance to either think about it or talk about it on their own or whenever there is a natural point for them to do something and me to get feedback on whether or not they understand or whether they have questions, I think that, that should be driving when I do my activities more than just timing but yeah so I think both of those things honestly play a role in when I decide to integrate it."* [*planned simultaneously with lecture content*]

Frank said, "*I try and focus on what I think is the most fundamental concept relevant to each of these topics and then how does that teach, how does that*

drive what else needs to be talked about with respect to that particular topic." [*planned after lecture is set*]

Active Learning Definition. There were three themes within this category: '*students engaged, involved, thinking about content', 'students interacting with each other',* and '*students as knowledge constructors'*. Instructors could be classified with more than one theme in this category. If they talked about students being engaged with the content, they were counted under that theme. If instructors instead or additionally talked about student collaboration as a part of what AL means it was counted as a separate theme. The last theme represented instructors talking about AL as a process where students construct their own knowledge. The way instructors defined or thought of AL also changed throughout the reform process (Table 3). In the first year most instructors (7) described AL as '*students engaged, involved, thinking about content'.* Only one instructor defined AL using the '*students as knowledge constructors'* theme. For example:

David said, *"Students are engaged in thinking about content rather than just remembering it."* [*students engaged, involved, thinking about content*]

Frank defined AL as, "*Students being engaged with the content and thinking about it in class."* [*students engaged, involved, thinking about content*]

Celine described AL as, *"Students building their own knowledge rather than being told facts."* [*students as knowledge constructors*]

In the second year, all instructors used the theme of '*students engaged, involved, thinking about content'* in their definition of AL. Some instructors (4) also began to add an interactive component to their definitions of AL: '*students interacting with each other'.* One instructor added the '*students as knowledge constructors'* theme to their definition of AL. For example:

> Bruce said, *"It means involvement, engagement, non-passive, doing things, thinking."* [*students engaged, involved, thinking about content*]
>
> David added an interactive component, *"That the students are engaged, that they're not just sitting and listening or taking notes but that they are required to interact with each other or interact with me or to actively be doing something you know evaluating data, coming up with a hypothesis, talking to each other, answering questions, something like that."* [*students engaged, involved, thinking about content* and *students interacting with each other*]
>
> Ingrid described AL as, *"So that means they need to do the learning themselves and not me. So basically, for me it's more the students have to do so they have to manipulate the concept, apply them to new situation, things like that."* [*students engaged, involved, thinking about content* and *students as knowledge constructors*]

And by the third year, instructors (7) defined AL as '*students engaged, involved, thinking about content'.* Five instructors added '*students interacting with each other'* and

three added '*students as knowledge constructors'* theme to their definition. For example:

> Albert described AL as, "*When the students are thinking about the material and doing something with it."* [*students engaged, involved, thinking about content*]
>
> Frank added an interactive component to his definition, *"It is when students are taking an active role in their own process of learning. They are working together to solve problems with the material."* [*students engaged, involved, thinking about content* and *students interacting with each other*]
>
> Erin added an interactive and construction component, *"Practicing!! Giving them an opportunity to practice the things I want them to be able to do on the exam, which then are also the things I think are important to learn in the class so anything that gets them to think. Consider, working together mainly just to think, to build their understanding."* [*students engaged, involved, thinking about content* and *students interacting with each other* and *students as knowledge constructors*]

Active Learning Type. There were five themes within the category of AL type: '*instructor used group AL',' instructor did not report using group AL', 'instructor used clickers', 'instructor used verbal questions',* and '*instructor used activities'.* The type of AL that instructors described implementing in their lectures changed across the three years (clickers, verbal questions, and activities) as well as the manner in which they asked students to work on those AL events (individually or in groups). In the first year,

most of the instructors reported that they (6) '*used group AL'* (Table 3). All instructors (8) '*used clickers',* most instructors (6) '*used verbal questions',* and some instructors (5) '*used activities'.*

> Erin used group AL, *"There's not a lot of individual, other than the clickers, well no actually cause I tell them to talk during clickers."* [*used clickers* and *used group AL*]
>
> When Harold was asked what role does student talk play in your classroom, he responded, "*Yeah…sorta zero."* [*instructor did not report using group AL*]
>
> Albert reported using clickers, "I use the clickers to review where we left off, assess preparedness, active learning, exams." [*used clickers*]
>
> Bruce reported using clickers and verbal questions along with group AL, *"Clickers, calling on students with verbal questions, small groups, and allow for students to share."* [*used group AL* and *used clickers* and *used verbal questions*]
>
> Celine reported using clickers, verbal questions, and activities, *"Open-ended questions that lead to discussion, think-pair-share, clicker questions, handouts"* [*used group AL* and *used clickers* and *used verbal questions* and *used activities*]

In the second year, all of the instructors (8) '*used group AL'* (Table 3). All instructors (8) '*used clickers',* almost all instructors (7) reported they '*used verbal questions',* and most instructors (6) '*used activities'.* The number of instructors that '*used verbal questions'* and '*used activities'* increased compared to year 1 (Table 3). For example:

Harold added group AL, "*Clickers, clickers with discussion. Just having students discuss something is something I've never ever done before and I've taught this four times."* [*used clickers* and *used group AL*]

Ingrid talked about verbal questions, "*I try whenever possible, try to find challenging questions that they can think about and I have them discuss it and we vote then discuss it, vote discuss, these kind of things*." [*used group AL* and *used verbal questions*]

Gail described adding activities to her classroom, "*This semester I've done three or four case studies where I have pulled from the Buffalo case study databank. Another activity we did a genetics worksheet where we did dominants and recessive, that was fun. We did an activity early on with DNA structure so students were given different components and had to orient a DNA molecule in the right orientation with the right components, um what else?"* [*used activities*]

By the third year all instructors (8) were implementing AL in a collaborative context (Table 3). Most instructors (7) still '*used clickers',* almost all instructors (7) '*used verbal questions',* and all instructors reported that they (8) '*used activities'.* One instructor reported not '*using clickers'* in year 3. The number of instructors '*using verbal questions'* and '*using activities'* increased over the three years (Table 3). For example:

Albert reported no longer using clickers, "*So this semester I didn't use clickers and frankly I didn't miss clickers and I don't have anything against them and I*

liked using them but this semester and I don't know how effective it was but I went with notecard approaches and group work on paper and they just turned that in and then I went over it and gave them participation points instead of actually grading it." [*used group AL* and *used activities*]

Celine commented on verbal questions, "*I know I ask a lot of questions and I try to make them, I think I have gotten better about this, I try to give students a chance to talk about it in groups or in pairs before I get responses from students."* [*used group AL* and *used verbal questions*]

Juanita talked about clickers, verbal questions and activities, "*Worksheets where they work together. Worksheets, they have to apply content from lecture notecards, questions from powerpoint to discuss, clickers with discussion, and I repoll if few correct."* [*used group AL* and *used clickers* and *used verbal questions* and *used activities*]

Teaching Style. There were three themes under teaching style where instructors reported whether the curriculum reform process impacted their teaching style: '*yes', 'no', 'do not know'.* In the first year, half (4) of the instructors did not see the reform affecting their teaching style or did not know how it would (Table 3). Those that reported '*no'* to their teaching style being impacted by the curriculum reform also indicated that it was because their teaching was already in a process of reform. The other instructors (4) reported that '*yes'*, the reform would affect their teaching style. For example:

Bruce revealed,*" I don't, my teaching style is always evolving regardless."* [*no*]

Frank stated, *“I don't really know. I'm sure it will but I already use active learning in the class.”* [*do not know*]

Gail said, *“I’ll have to be more organized and less flexible, especially when we have the TA’s lead the discussion section for one class each week. I will have less time to get the information to the students and will need to stick to my schedule.”* [*yes*]

The second year only one instructor still thought that curriculum reform would not affect their teaching style and almost all instructors (7) said that the reform would affect their teaching style. For example:

Juanita said, *“Not really because that’s you know that’s the whole, I don’t think it will change. I think it’s just going to change what percent is done outside of the class and what percent is done in the class.”* [*no*]

When David was asked he said, *“Well, just the shift in thinking of how the material is presented so focusing not only the students but my attention on the fact that these are connected concepts rather than individual bite-sized pieces of information and so like I said the same basic content but put into a different context. It can be very different and I think it is much more effective to think about, teach about it in that connected context that it is discrete pieces of information.”* [*yes*]

The third year interviews recorded most instructors (7) commented on how the reform had affected their use of AL, while one reported the reform did not impact his teaching style and that lack of change was due to their teaching changing prior to reform. For example:

> Bruce still did not see the reform impacting his teaching style, "*Not really. My teaching style has been evolving in this way already."* [*no*]
>
> Harold stated, *"It's basically hauled me out of my safety zone. Cause I like to lecture. I have fun. I've started using activities. I guess one other thing I've gotten better at is asking questions of the class and not answering it five seconds later myself."* [*yes*]
>
> Ingrid said,*" I think you know if we talk about awhile ago I was not doing any of this, learning, active learning. I was not thinking too much about it. I think a lot more about what I would like, I'm more purposeful in the way I design my class. I'm thinking more about what do I want the students to know so this backward design, I'm doing this a lot more."* [*yes*]

Assessment. Unlike the other categories identified, assessment was a category that was identified a posteriori and emerged from the data itself. After noticing instructors mentioning assessment in their interviews, all responses related to assessment were marked and re-read. Assessment did not have any themes. When instructors mentioned assessment, it was always their reflections about what to test, how to test, and using assessment itself as an instructional strategy. The frequency of

instructors that talked about assessment in their interviews increased over time. In the first year, only one instructor talked about assessment during their interview. For example:

> When asked what he would like to change about his course in the future Albert said, "*I never did review before an exam but I think I should. I would like to add a review before exams and have another form of assessment, more evaluator-y quizzes.*

During the second year, two instructors talked about assessment in terms of aligning exams with the activities that are done in the classrooms and one talked about assessment from a classroom strategy standpoint. For example:

> When asked what was not working in her course Erin answered, "*I'm still struggling with assessment in the course. I'm happier with my assessment this year.* [My TA] *helped me to revise how I ask question so that they are scenario based and so now I think that the assessment is much more in line with the activities that we do in class."*
>
> Ingrid also talked about aligning assessment with classroom practices, "*I try to give them group activities where I have them apply concepts and I do in groups because I think it's too hard to do it on your own if you've never done it. So I think that the first thing and then I always test them in that on an exam so it's not like I have them do it and then forget about it. And they know it so they*

always apply to this."

When asked about practices that help his students to learn Bruce said, "*Really my exams are themselves a strategy. My first exam is a take-home. My second exam is in-class and then take-home, same exam, and then their score is the average and then the third exam it's time to now fly by yourself. I've always had a part of the exam that are questions that I project and they can talk amongst themselves.*"

During the third year of reform, six instructors mentioned assessment including the instructors that had begun to think about assessment in previous years (Albert, Erin, Ingrid and Bruce) and an additional two: Celine and Frank. For example:

Celine talked about using assessment in terms of figuring out what is less important to cover with the reduced lecture time, "*I try to pick out things that I talked about and taught students in the past that I didn't end up testing them over very much cause I figured it's something that's not important enough for me to test them over or if it's really like difficult for me to test them on then that seems like an obvious thing to take out of the course."*

Frank said, "*I give a lot of analysis and interpretation and the students realize this. And I say look, I want you to be able to think this stuff but it is hard*."

Discussion

Instructor beliefs about teaching and learning have been hypothesized to shape the practices they actually implement in the classroom (Pajares, 1992). However, it has been shown that what instructors say about their teaching may not match with what actually occurs (Ebert-May et al., 2011), revealing a disconnect between the perception and reality of instruction. This study provides evidence that changes in beliefs can shape practices over time when instructors are engaged in a curriculum reform and suggests that for this to happen, instructors must (1) engage in reflection on what constitutes best practices and their own use of those practices, (2) their reflection is facilitated by a collaborative component, and (3) time is needed.

Instructor AL use significantly increased across the three years of curriculum reform, although changes between years were not significantly different. This may suggest that meaningfully reforming instruction can be accomplished by making small programmatic changes that accumulate over time to create the change, thus reducing barriers such as time for planning (Dancy & Henderson, 2010). Indeed, the average use of AL in the lecture classes increased by 12% from the first year to the third year of reform, despite the loss of 1/3 of the class time. This suggests that if you embed the changes in practices into the program, they can persist despite the addition of new barriers such as loss of class time.

When lecture was paired with discussion in year three, AL overall increased significantly as did group AL, the use of verbal questions, and the use of activities. Clearly, a change in the format of introductory courses to include one small group meeting per week can have additional significant impacts on AL use that may be more

difficult to achieve in large class settings alone. While this increase in AL by shifting instructional time from large class to small class environments is not surprising, institutions undergoing reform may consider innovative ways to reallocate class time and instructors to achieve the types of instructional environments that have been shown to achieve more success for all students, but particularly for underrepresented students (Freeman et al., 2014).

One limitation of this study was that the quality of the AL that occurred in the classroom was not assessed. This will be a necessary next step to better understand the conditions under which AL implementation can improve student outcomes (Andrews, Leonard, Colgrove, & Kalinowski, 2011). Another limitation was the number of observations. The specific AL techniques used by an instructor can fluctuate over different parts of the course, so three classroom visits may not accurately capture the true nature of AL use in each class. However, three observations have been used in similar observational studies (Eddy, Converse, & Wenderoth, 2015).

Although faculty were able to increase their use of AL and maintain this increase when losing instructional time, the first two years of this study showed extremely similar and sometimes identical numbers. The shifts in use of AL started slowly and then gained momentum in the third year. This may be due to instructor need for time to think about and process change. In the same way that students need time to process and construct new information, processing time may be an important factor in achieving pedagogical shifts in faculty instruction. Part of this may be because shifts in practices require shifts in how instructors define AL, and much like misconceptions, it may be hard to dislodge these beliefs. Having a supportive group may be instrumental in

persevering with change, as faculty may want to abandon efforts when change does not happen instantly (Henderson & Darcy, 2007).

The interviews provided evidence for how instructional change among the faculty was achieved. There were shifts over the course of the reform in how faculty defined, planned for, and implemented AL in their classrooms. As faculty began to talk about using learning objectives to align their course, this in turn seemed to foster a greater programmatic focus on assessment. Faculty also changed how they conceptualized AL across the years. Almost all instructors saw AL as engaging or involving the students with the content in the first year of reform. During the second year, instructors built onto that definition by saying that AL is interactive and students work together. This was reflected in their classrooms as all instructors reported using and were observed to use group AL that year. In the third year, instructors also started to see AL as a process where students are constructors of knowledge. This may have been why more instructors self-reported using activities in their classes in year 3, because they typically provide more extended opportunities for group work and knowledge construction.

Part of the shift in faculty practices and beliefs may be attributed to the use of feedback (Kluger & DeNisi, 1996) and collaboration (Sunal et al., 2001) throughout the reform process, factors that have been shown to assist faculty when dealing with barriers to change (Gormally, Evans, & Brickman, 2014). The AL observation results were shared at the end of each semester so that each instructor was aware of their own use of AL and the use for the program as a whole. Instructors then used this information as part of their discussions about programmatic goals for AL; increasing the use of group AL, for example, was discussed as a change the group wanted to make over

time. Through this feedback and discussion, instructors were also introduced to the variety of ways AL was being implemented in the classroom when they shared ideas with each other. As such, it is recommended that institutions fostering reform should host regular meetings with the instructors where they are asked to discuss and make decisions about implementation of their courses; this may be key to the long-needed switch from advocating for AL to implementing AL (Freeman et al., 2014; Gormally, Evans, & Brickman, 2014).

The monthly meetings instructors participated in included discussion of best practices, learning objectives, backward design (Wiggins & McTighe, 1998), using collaboration in classroom activities, as well as professional development sessions. Several instructors commented during the interviews on the value of these meetings and the collaborative community context in which they occurred. Frank said, *"It's allowed me to interact with the other instructors of the course and see what kinds of activities they do."* Celine talked about the impact on her teaching, *"I think having the data on what I was doing in my classroom presented to me and then being able to compare myself with other people that are teaching the same course, I think it led to me using more, asking students to talk to one another more and I think I got ideas for activities that I may not have otherwise done."* Ingrid also shared about the impact in her classroom, "*I think a lot more about what I would like, I'm more purposeful in the way I design my class. I'm thinking more about what do I want the students to know so this backward design, I'm doing this a lot more."*

Instructional change to increase the use of AL in large introductory biology courses is possible, though this study suggests that it may require time for the changes

that have been discussed to take effect. A recommendation for departments that wish to align their instruction with best practices is to do it collaboratively and to give faculty some autonomy to make decisions about implementation. Faculty should meet regularly and talk about their classroom practices, share ideas and resources, discuss relevant literature, and conduct structured observations of their practice in order to provide the type of feedback necessary to change practice. Through community approaches such as these, introductory biology may move closer to the types of practices that can help all students achieve success.

Appendix

Instructor Interview Questions

1. Will you give me a brief history of your training and experience related to teaching?

2. What courses are you currently teaching?

3. What are you goals for the course?

4. Will you walk me through the process of designing an average class within your course?

5. What does active learning mean to you?

6. When you are designing a class for your course, do you think about how to integrate active learning?

7. What types of active learning do you use in your course?

8. What would you estimate is the percentage of class time you spend in active learning activities?

9. Besides active learning, what other strategies do you use to help your students learn?

10. Tell me about your perception of the course, what is working, what isn't?

11. What would you like to change about your course in the future?

12. What are your thoughts on curriculum reform?

13. How do you see curriculum reform affecting your teaching style?

14. Do you have any other thoughts on the course you would like to share?

CHAPTER III

CURRICULUM ALIGNMENT WITH VISION AND CHANGE IMPROVES STUDENT SCIENTIFIC LITERACY

Abstract

The University of Tennessee, Division of Biology, recently implemented curriculum changes to the majors' two-course introductory biology sequence in alignment with the concepts and competencies described in the *Vision and Change* final report (AAAS, 2011). Discussion sections were added to both lab and lecture portions of the sequence and lab time associated with the courses was reduced. The discussion sections were backward designed using learning objectives with the goal to develop 'process of science' skills. The curriculum was intended to be implemented using active learning techniques paired with student collaboration. This study sought to determine if these changes resulted in a higher gain of scientific literacy for students by assessing two cohorts of students using the Test of Scientific Literacy Skills (TOSLS: Gormally, Brickman, & Lutz, 2011) in a pre-post design. Retention of student scientific literacy skills was also assessed. Overall gain in scientific literacy was found to be two times greater for the students experiencing the reformed curriculum. Retention of scientific literacy skills did not differ between the cohorts overall, however, the cohort experiencing the discussion sections and reduced lab time did have a higher gain and retention of scientific literacy skills related to research design.

Introduction

Scientific literacy can be described as the knowledge and understanding of scientific concepts and processes (AAAS, 1989). Student understanding of scientific literacy has been advocated as a necessary outcome of an undergraduate science degree in order to produce a workforce capable of the technology comprehension and national competitiveness needed to handle the social, economic, and environmental challenges

of today's society (AAAS, 2011). Given this, enhancing the scientific literacy skills of science majors should be a major focus in science curricula. Core scientific literacy skills can include identifying hypotheses, understanding experimental design, reading figures, finding evidence for claims, and creating scientific models. The *Vision and Change in Undergraduate Biology Education* final report (AAAS, 2011) recommends restructuring introductory biology curricula in a way that emphasizes core biological concepts *and* competencies and implementing these changes using student-centered (active learning) practices.

Although there have been many calls for the inclusion of scientific literacy skills in science curricula (AAAS, 2011; NRC, 2004), studies that have reported how these curricula impact student proficiencies in scientific literacy are lacking. The short-term acquisition of scientific literacy is important for the biology graduates of today, but more importantly is the retention of these skills, as they are needed after graduation when students start their careers. While there are relatively few studies that have assessed student scientific literacy in students (Dirks & Cunningham, 2006; Gormally, Brickman, & Lutz, 2012; Porter et al., 2010), even fewer (Khishfe, 2015) have looked at the retention of such skills despite the agreement among policy makers, educators, and researchers that these skills are an important long-term goal of science education. Retention studies are difficult, however, because of the need to re-test the same students and disagreement about what constitutes a reasonable length of time to measure retention. Long-term retention of information and course material in undergraduates has been assessed, for example, 21 days following an intervention

(O'Day, 2007), four months later (Crossgrove & Curran, 2008; Khishfe 2015), and as long as three years later (Derting & Ebert-May, 2010).

Many studies on scientific literacy have focused on student analysis of primary literature and scientific writing tasks. Kozeracki et al. (2006) found that a primary literature program benefited student scientific literacy. In a study examining the enhancement of scientific writing skills, it was found that the greatest predictor of success was prior experience in scientific writing tasks (Jerde & Taper, 2004). Porter et al. (2010) identified gains in information literacy related in students' ability to access, retrieve, analyze, and evaluate primary scientific literature. Gormally, Brickman, & Lutz (2012) showed that students experiencing a project-based non-majors' biology course had higher gains in scientific literacy compared to a traditional non-majors' course. However, retention of these literacy skills was not investigated. Khishfe (2015) investigated retention as it relates to nature of science and found that four months after participation in a unit about genetic engineering, many students reverted to misconceptions regarding the topic.

Context of the Study

This study took place at the University of Tennessee while undergoing a curriculum reform within the science majors' introductory biology sequence. The sequence was composed of two courses: one that focused on topics of cell biology and the other on organismal biology. The curriculum reform was focused on transforming these two courses to bring them into alignment with the recommendations of the *Vision and Change* final report (AAAS, 2011). Prior to the reform, the courses were each taken for 4-credit hours, had large lectures that met for 150 minutes each week, and had an

associated 3-hour weekly guided inquiry lab (Buck, Bretz, & Towns, 2008) taught by graduate teaching assistants (GTAs). After the reform, the labs were separated from the lectures to form their own independent course worth 2-credit hours (3-hour lab and a science process skills discussion taken once during the introductory sequence). The labs remained in the same format (guided inquiry), using half of the same exercises featured in the previous two semester curricula. At the same time, discussions to explore each exercise more fully were added to the course. The two lecture courses were reduced to 3-credit hours each and met in large lectures for 100 minutes per week and in small-group GTA-led scientific literacy discussions for 50 minutes per week. These scientific literacy discussions were specifically designed to build core scientific literacy competencies such as identifying and creating hypotheses and experimental design, interpreting figures and data, using quantitative reasoning and modeling, and analyzing scientific arguments. The discussions also explicitly used small-group active learning pedagogy as recommended by the *Vision and Change* final report (AAAS, 2011).

The discussion curricula were created using backwards design (Wiggins & McTighe, 1998) where the intended outcomes are used to plan instruction by community members who represented constituents of the course (NSF DUE 1245215; PI Schussler). Members of the community included faculty, postdocs, graduate, and undergraduate students from the three biology departments served by the courses. The community worked in small groups over the 2013 – 2014 school year to plan the new curricula around specific process of science learning objectives and utilizing primary literature. Many researchers call for the inclusion of information literacy skills in order to

achieve scientific literacy skills (Glynn & Muth, 1994; Firooznia & Andreadis, 2006) and others cite using primary literature as a way to build scientific thinking (Gottesman & Hoskins, 2013). The curricula used a scaffolding process in which skills were progressively learned by students and then built upon throughout the courses.

Reformed curricula activities were planned to require student collaboration as a way to encourage active learning. A meta-analysis conducted by Freeman et al. (2014) found that active learning, where students are explicitly asked to engage in thinking about course material during class, increases student success for students in engineering, science, and mathematics. There is some evidence that active learning may also result in better retention of understanding than traditional lecture, but these studies require a time delay in the measurement of student outcomes (Derting & Ebert-May, 2010; Dougherty, et al., 1995). Assessing the gain and retention of scientific literacy skills as a result of curriculum changes is important in order to evaluate the effectiveness of this *Vision and Change* reform to foster new competencies in students at the University of Tennessee. In order to adequately compare the previous to the reformed curricula, this study compared the gains in and retention of scientific literacy skills for students in the 2013 – 2014 introductory sequence who experienced two semesters of lab and did not experience scientific literacy discussions (Cohort 1) with students from the 2014 – 2015 cohort who experienced one semester of lab and three scientific literacy discussions (Cohort 2). This study addressed the following questions:

1. **Does the curriculum alignment with Vision and Change competencies result in a higher gain of scientific literacy skills for students?**

2. Does the curriculum alignment with Vision and Change competencies result in a higher retention of scientific literacy skills for students?

Methods

Participants

The participants were undergraduate students at the University of Tennessee registered for all courses of either the majors' introductory biology sequence in the 2013 – 2014 or 2014 – 2015 academic years. A total of 305 participants were included in this study. The first cohort (N_1) had 156 matched participants and the second cohort (N_2) had 149 matched participants; of these, 20 participated in the retention measure for N_1 and 50 for the retention measure for N_2. All students present on the days the assessments were given were potential participants, but only those who took both the pretest at the beginning of the first course (fall 2013 or 2014) and the posttest at the end of the second course (spring 2014 or 2015) for a given academic year were included in the results Student participation in the assessments was voluntary, and the Institutional Revenue Board approved the research prior to data collection.

Instrument

The Test of Scientific Literacy Skills (TOSLS; Gormally, Brickman, & Lutz, 2012) is a fixed choice (multiple choice) instrument with one defined correct answer for each of the 28 questions on different aspects of scientific literacy. The TOSLS assesses nine skills that all contribute to the construct of scientific literacy and has a reliability estimate above 0.70 using the Kuder-Richardson Formula 20 (KR-20). Due to class time constraints for the assessment associated with this study (20 minutes), only a portion of the TOSLS was used. The original 28 question test was designed to be implemented in

35 minutes. In keeping with the same average time allowance per question, 16 items aligned with 7 skills were chosen for this study (Appendix). These questions were selected by letting the faculty associated with the courses choose the TOSLS skills that best represented the competencies outlined in the *Vision and Change* final report (AAAS, 2011). These skills (each with associated TOSLS question(s)) included: identify a valid scientific argument, understand elements of research design, make a graph, interpret a graph, solve problems using quantitative skills, understand and interpret basic statistics, and justify conclusions based on quantitative data. Student scores on the TOSLS assessment were calculated by summing the number of correct survey item responses, thus student scores ranged from 0 – 16 (0% – 100%).

Data Collection

Both cohorts were assessed for scientific literacy at three different time points: the beginning of the fall semester, the end of the spring semester, and (for retention) the fall semester following completion of the sequence (Table 6). Cohort one completed this assessment schedule in 2013 and 2014, and cohort two completed it in 2014 and 2015. The pre-test was given to the first cohort during their first lab class period in the first introductory course while the second cohort took theirs within their first discussion class period of the first introductory course. The post-assessment was given to cohort one during the last lab class period of the second introductory course and the second cohort took the post-assessment during their last discussion class period of their second introductory course. All administrations of the TOSLS were performed at the beginning of class. The graduate teaching assistants (GTAs) teaching the labs or discussions administered the pre- and post-assessments by using a scripted introduction and set of

Table 6. Schedule of three TOSLS implementations for each cohort.

Group	Fall 2013	Spring 2014	Fall 2014	Spring 2014	Fall 2015
Cohort One (2 labs 0 Discussions)	Pre (1)	Post (2)	Retention (3)		
Cohort Two (1 lab 3 Discussions)			Pre (1)	Post (2)	Retention (3)

instructions for implementation. Students filled out a demographic sheet after the completion of the TOSLS to prevent stereotype threat (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). The survey instructions were presented by the first author at the GTA preparation meeting the week before implementation. All student assessments were sealed in an envelope by the GTA and immediately collected.

This study featured a mid-length retention testing time of four months to ensure enough time for retention, but to prevent loss of participants from the research project. The sophomore level Genetics, Microbiology, and General Ecology courses were chosen as the target courses in which to implement the retention TOSLS because these are the 200-level courses that biology majors would progress into after completing the introductory sequence. Retention tests were given to students by the first author during the first lab and discussion periods of each of the three classes in fall 2014. Completed assessments were sorted to identify students who had taken both the pre- and post-assessments as part of cohort one. This method produced a very low number of matches (20 out of the 156 matched pre-post for cohort one) because many students in the majors' introductory sequence are pre-professional students who diverge from the typical biology sequence in their sophomore year. To increase the retention sample for the second cohort, a different recruitment approach was used in fall 2015. All students who had completed the pre- and post-assessments for cohort two ($N_2 = 149$) were emailed and invited to participate in an assessment of scientific literacy for the Division of Biology. All participating students were given a $10 credit on their student account for completion of the survey. Although this method resulted in more retention participants

(N_2 = 50), it still did not produce a retention measure for the majority (66%) of the sample.

Data Analysis

Student pre- and post-assessments were matched using student identification numbers and not names. All non-matching data (e.g., a student who took the pre- but not the post-assessment) were excluded from the analysis. Demographics for students in the sample were matched to their assessments and recorded in a spreadsheet. Finally, to make sure the comparison encompassed the full introductory sequence each year (lectures plus lab(s)), students in the second cohort who did not self-report taking the single semester lab course in fall 2014 or spring 2015 were also excluded from the final sample (N = 34). Student retention scores were matched to pre-post score data using student identification numbers and not names. Eight participants were eliminated from the second cohort's retention sample because they completed the survey in less than five minutes, which was judged to be insufficient time to adequately read and complete the assessment.

The pre-score data and post-score data were also divided by skill. Student scores were used to calculate an average percent of responses correct for the specific skill by averaging performance on the questions assigned to each of the seven skills. For example, read and interpret graphical information has four questions and if a student were to get 3 out of 4 of those questions correct, their score for that skill would be 0.75. Two sets of data were used for the analyses of gain and two for retention: pre-post assessment scores overall and by skill, and post-retention assessment scores overall and by skill.

The data were then checked to ensure all values were within the specified range of 0-16, assessed for outliers, and assessed for normality. Outliers were modified to be one unit from /3/ standard deviations (SD) from the mean (Tabachnick & Fidell, 2001). After adjusting the outliers, a Shapiro-Wilk test was used to assess each variable's normality. Non-normal variables resulted in the use of a more stringent alpha level (p = 0.025) rather than transforming the variables in order to maintain the students' true scores.

The first research question was to identify whether students experiencing the new curriculum gained in scientific literacy as compared to students who experienced the old curriculum. Using the data of pre-post assessment scores for each cohort, a one-way mixed analysis of variance (ANOVA) with the pre-score and post-score as the dependent variables and cohort as the independent variable was calculated to determine if there was an interaction effect between the gain in scores over time and the curriculum experienced. Independent t-tests were also conducted with pre-score as the dependent variable and cohort as the independent variable to determine if the two cohorts began their courses in similar places with regards to scientific literacy skill set. An analysis of covariance (ANCOVA) was performed to compare post-test scores while controlling for different pre-tests scores to separate a possible cohort effect from curriculum effect. All analyses were conducted using SPSS (IBM, 2011).

Student scores were also compared using student demographics to ensure equity among students within cohorts. Due to low numbers within the category of ethnicity, the variables were reduced to white and non-white students. Two separate mixed ANOVA with gender (male/female) and ethnicity (white/non-white) as the

independent variables and pre-score and post-score as the dependent variables were calculated. A repeated measures ANOVA was used to determine if there were differences in pre-score and post-score for students by year in school at the university

To determine if there were differences in gain by skill between the cohorts, separate one-way mixed ANOVAs with the pre-score and post-score of each skill as the dependent variables and cohort as the independent variable was calculated. For significant results, independent t-tests were conducted with pre-score as the dependent variable and cohort as the independent variable to determine if the two cohorts began their courses in similar places with regards to scientific literacy skill. An analysis of covariance (ANCOVA) was performed to compare post-test scores while controlling for different pre-test scores to separate a possible cohort effect from curriculum effect. All analyses were conducted using SPSS (IBM, 2011).

The second research question was to identify whether students experiencing the new curriculum had improved retention of scientific literacy as compared to students who experienced the old curriculum. Not all of the students from either cohort participated in the retention measure. Data analyses were performed using the students that did participate. A mixed ANOVA with post-score and retention score as the dependent variables and cohort as the independent variable was calculated. To determine if there was a difference in retention of particular skills between the cohorts, a repeated measures ANOVAs was used to compare cohorts by specific skill. For significant results, independent t-tests were also conducted to determine if the two cohorts had similar post-scores with respect to specific skill. All analyses were conducted using SPSS (IBM, 2011).

Two separate repeated-measures ANOVAs with gender (male/female) and ethnicity (white/non-white) as the independent variables with post-score and retention as the dependent variables was calculated. A repeated-measures ANOVA was used to determine if there were differences in post-score and retention for students by year in school at university.

Results

The first cohort showed a gain of 7% in scientific literacy scores from pre-test (M = 0.64, SD = 0.17) to post-test (M = 0.71, SD = 0.17). The second cohort showed a gain of 13% from pre-test (M = 0.58, SD = 0.17) to post-test (M = 0.71, SD = 0.71). The cohorts' retention scores were similar in score to each other (M_1 = 0.70, SD_1 = 0.14; M_2 = 0.71, SD_2 = 0.21).

Not all students who took the pretest (N_1 = 610; N_2 = 602) also took the posttest (N_1 = 490; N_2 = 394). However, both cohorts had similar numbers of students with matched pre- and post-assessments (N_1 = 156; N_2 = 149). Of these matched students, both cohorts had low numbers of students who also completed the retention test (N_1 = 20; N_2 = 50). This sample included 18.8% identifying as ethnic/racial minorities, which is representative of the overall student body of the institution (19%), and 56% identifying as female which is slightly higher than the overall student population (49%). The two cohorts did not differ significantly in terms of gender, ethnicity, or year in school. Most (59.1%) students were first-year freshman. There were also sophomores (24.7%), juniors (11.1%), and seniors (3%) in the sample.

Scientific Literacy Gain

There was an overall significant interaction effect when comparing the cohorts across time, with cohort 2 showing a higher gain in scientific literacy pre-test to post-test compared to cohort 1, $F(1, 303) = 8.373$, $p = 0.004$, $\eta^2 = 0.027$, power = 0.822. An independent samples t-test confirmed that pre-test scores for cohort 1 ($M = 0.64$, $SD = 0.17$) were significantly higher when compared to cohort 2 ($M = 0.58$, $SD = 0.17$), $t(303) = 2.771$, $p = 0.006$, and that there was no difference in post-test scores. The cohorts had different averages on the pre-test but similar averages on the post-test, meaning the cohort 2 had greater gains in scientific literacy when compared to cohort 1 (Figure 4). An ANCOVA controlling for the differences in pre-scores between the cohorts revealed only marginal significance in post-scores ($p = .08$). There were no differences in gain of scientific literacy skills for students when comparing gender and ethnicity. There were also no differences in gain among students in different years in school.

Separate repeated measures ANOVAs were used to determine if the cohorts differed by gain in specific skill areas and revealed significant differences between cohorts on two skills: identify a valid scientific argument ($F(1, 303) = 5.209$, $p = 0.023$, $\eta^2 = 0.017$, power = 0.624) and understand elements of research design ($F(1, 303) = 77.888$, $p = 0.000$, $\eta^2 = 0.204$, power = 1.0). The first cohort's pre-score ($M = 0.83$, $SD = 0.38$) was similar to the second cohort (M = 0.87, $SD = 0.34$) in the skill of identify a valid scientific argument. Post-scores were also similar. However, the first cohort ($M = 0.91$, $SD = 0.29$) scored higher on the post-test compared to their pre-scores while the second cohort ($M = 0.84$, $SD = 0.37$) scored lower in the post-test than their pre-test (Figure 5). The opposite was true for the skill of understand elements of research

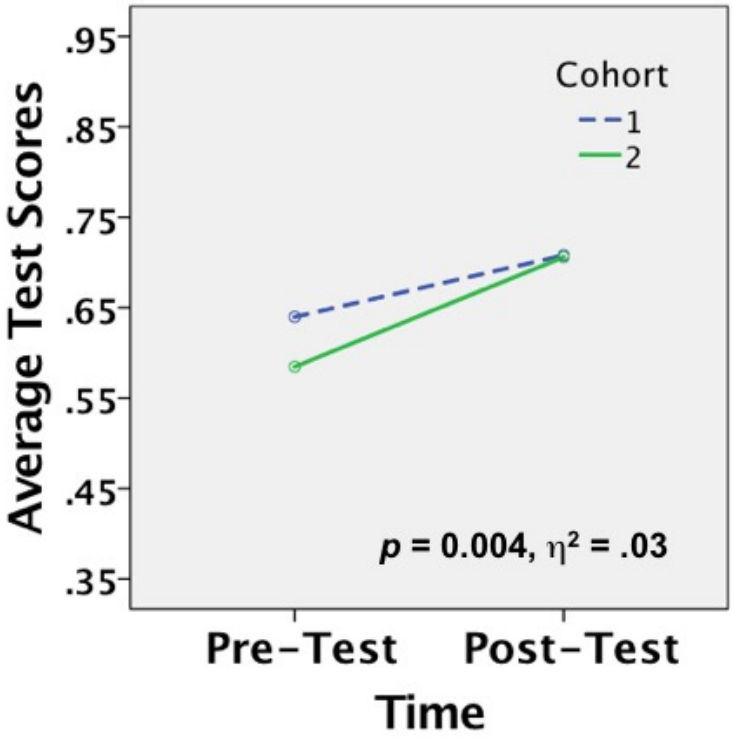


Figure 4. Overall gain from pre-test to post-test in scientific literacy for cohort 1 and cohort 2.

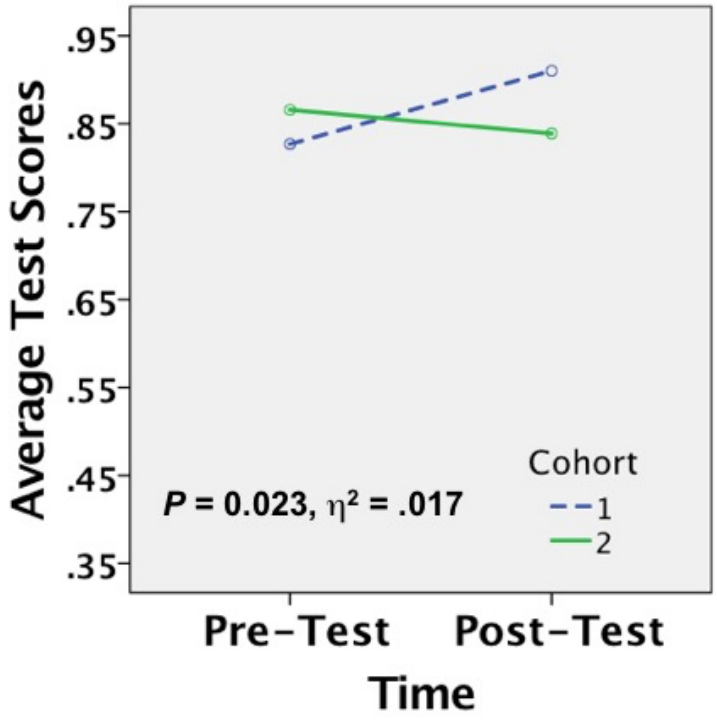


Figure 5. Gain from pre-test to post-test in the skill of identify a valid argument for cohort 1 and cohort 2.

design. For this skill, the second cohort (*M* = 0.38, *SD* = 0.27) had significantly lower pre-scores when compared to the first cohort (*M* = 0.72, *SD* = 0.33), *t*(303) = 9.997, *p* = 0.000. By the post-test, the second cohort's scores (*M* = 0.78, *SD* = 0.74) were similar to the first cohort (*M* = 0.74, *SD* = 0.35; Figure 6). An ANCOVA controlling for the difference in pre-test scores on this skill still showed a difference in post-scores, F(1, 302) = 13.047, p = 0.000, η^2 = 0.041, power = 0.95. Although not significant using our more stringent alpha level of 0.025, two skills revealed marginal significance: read and interpret graphical representations of data (*p* = 0.029) and understand and interpret basic statistics (*p* = 0.031). When comparing gain on the other skills by cohort, no significant results were found: make a graph (*p* = 0.43), solve problems using quantitative skills (*p* = 0.23), and justify inferences, predictions, and conclusions based on quantitative data (*p* = 0.84).

Scientific Literacy Retention

There were no differences in the retention of scientific literacy scores by cohort (Figure 7). There were also no differences in retention of scientific literacy for students when comparing gender and ethnicity or between students in different years. When breaking down retention by skill, differences were found for understand elements of research design, *F*(1, 68) = 5.954, *p* = 0.017, η^2 = 0.081, power = 0.672 (Figure 8). Independent t-tests revealed that the cohorts had similar scores in this skill on the post-test, but the second cohort (M = 0.72, SD = 0.37) scored significantly higher on the retention measure when compared to the first cohort (M = 0.53, SD = 0.11), *t*(65.335) = -3.389, *p* = 0.001. When comparing retention on the other skills by cohort, no significant results were found: identify a valid argument (*p* = 0.67), make a graph (*p* = 0.53), read

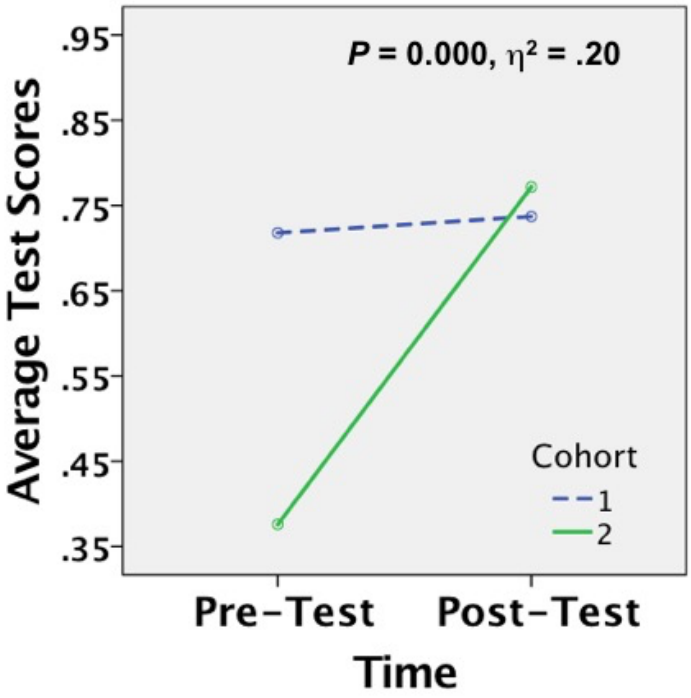


Figure 6. Gain from pre-test to post-test in the skill of research design for cohort 1 and cohort 2.

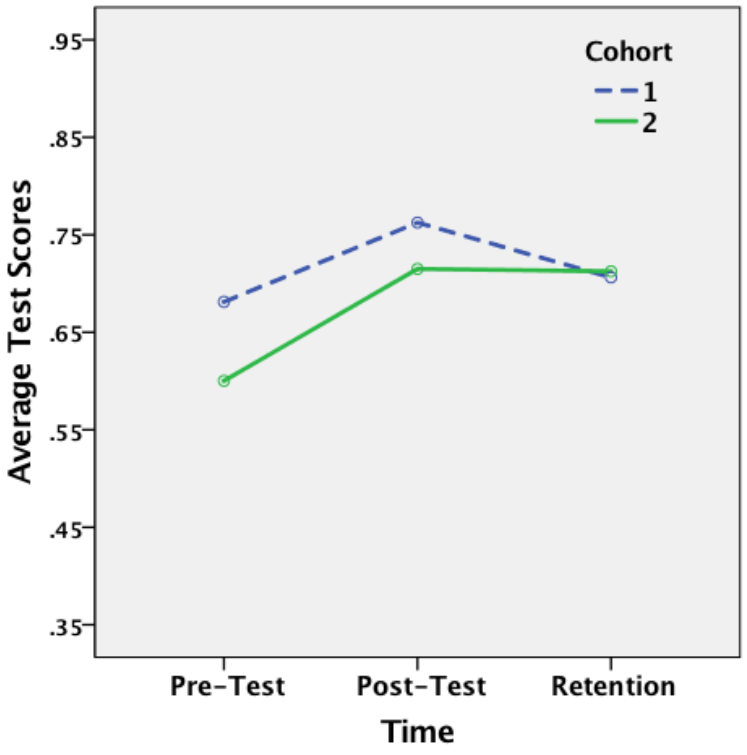


Figure 7. Retention sample pre, post, and retention test scores by cohort.

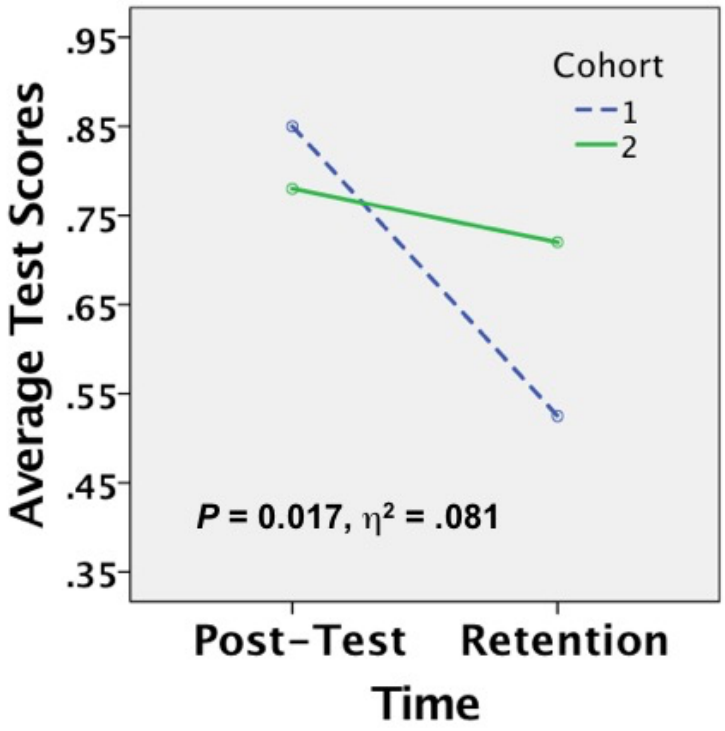


Figure 8. Retention from post-test to retention measure in the skill of identify research design for cohort 1 and cohort 2.

and interpret graphical representations of data (p = 0.26), solve problems using quantitative skills (p = 1.0), understand and interpret basic statistics (p = 0.59), and justify inferences, predictions, and conclusions based on quantitative data (p = 0.73).

Discussion

Previous studies (Jerde & Taper, 2004; Kozeracki *et al.*, 2006) have shown that when specific educational outcomes are desired, curricula designed to explicitly target those outcomes can be successful. Porter et al. (2010) found that students improved their information and scientific literacy skills after participating in exercises targeting those specific skills. Our results have added to this, showing that discussion curricula designed to foster scientific literacy skills in introductory courses for biology majors can increase aspects of scientific literacy and that these effects are retained as long as four months after students have completed the courses. This study has also provided evidence for the effectiveness of curricula aligned with the competencies described within the *Vision and Change* final report (AAAS, 2011) to improve scientific literacy, and that these improvements do not necessarily need to be achieved by increasing laboratory experiences, but can instead be achieved through active learning exercises centered on scientific literature. Intentionally designed curricula to enhance a particular set of skills can indeed make a difference in introductory biology student outcomes.

While students in both cohorts experienced gains in scientific literacy over the academic year, greater overall gains were achieved by the students who experienced the curriculum designed to align with the *Vision and Change* competencies (AAAS, 2011). In particular, the second cohort showed a greater gain related to research design

than the first cohort, which appeared to play a large role in the overall literacy score differences. The first semester lecture discussion has a focus on diagramming the research designs of the primary literature papers that students read as part of their activities. These diagrams are then explicitly aligned with the hypothesis and graphs of results shown in the scientific paper being read. It may be that this visualization of the research design is an important step to foster student understanding of this critical skill (Ainsworth, Prain, & Tytler, 2011). The discussion cohort also demonstrated a greater retention of research design skills four months after the course (and for most students there was an eight-month gap from the first semester discussion to the retention test date) suggesting that the practices that enhanced this skill were very effective in embedding this understanding into students' minds. It is interesting to note that a skill such as research design that would be expected to arise from laboratory study was enhanced even in the context of decreased lab time. It may be that the study of primary literature, when done correctly, may be more beneficial to the development of this skill to introductory students than designing experiments in lab may be.

The ability to identify a valid scientific argument was chosen by faculty as an important part of student scientific literacy, yet the discussion cohort displayed a slight but not significant decrease in that ability after experiencing the new curriculum. This was particularly surprising since the second lecture discussion focuses almost exclusively on scientific argumentation as a central learning objective. However, this skill was measured via only one question on the TOSLS and it may be that the manner in which the skill was presented was not similar to the way it was addressed in the discussions. It should also be noted that the differences between cohorts in this skill

disappeared by the retention time point, indicating an equal understanding four months later.

Both cohorts were equally able to retain their scientific literacy gains as measured four months after the completion of the courses. Since the reformed curriculum students experienced greater gains in scientific literacy, their scientific literacy would be retained at this higher level. This indicates, however, that there is no evidence that the active learning discussions employed in this study improve scientific literacy retention compared to other curricula previously employed. On the other hand, it is encouraging to know that students retain a good amount of the scientific literacy that they obtain from their introductory coursework. Scientific literacy is an important skill to foster within science graduates as it is not only used for scientific careers, but also in real-world every day scenarios that require evaluating data and using it to make informal decisions (AAAS, 2011), therefore student retention of what they learn in school is important.

While these results are encouraging to those considering a *Vision and Change* curriculum reform, there are some limitations. Confounding effects may exist between the addition of discussions and reduction in labs and there is no way to ensure that it was the addition of discussions alone that increased scientific literacy. There was also no way to control for other science courses students were taking over the academic year that may have impacted the results. In addition, the low numbers of participants within the retention sample make it hard to generalize these results to the students within our sample nor to all biology major students. There were also differences in the methods used to obtain the retention data and it should be noted that there are likely

resulting differences in the student retention sample as far as motivation or other factors. Students who took a test as part of the class work in a sophomore level biology course may be very different from students that are given an incentive to take a test. Therefore, there may be a higher or lower retention of scientific literacy for students in the discussion sections that this study was unable to capture. Finally, the TOSLS instrument was not able to be used in its entire validated form, which may have impacted this measure of scientific literacy. For each administration of the TOSLS, a KR-20 was calculated and it was found to range from 0.54 – 0.76. A score of 0.70 or above is considered acceptable (Cronbach, 1951). The original instrument reported a reliability of 0.731 and 0.748. The discrepancy of reliability is likely due to the small numbers of participants in the retention measures and not using the instrument in its entirety.

Despite these limitations, this study provides a unique snapshot of how the revision of curricula to align with *Vision and Change* may impact undergraduate biology majors. Assessing the effectiveness of curriculum changes such as these is necessary to ensure that students are achieving the intended outcomes of a course or set of courses (Resnick & Resnick, 1992). These data indicate that students do in fact acquire scientific literacy early in their undergraduate careers, that these skills can be enhanced, and that they will retain these skills over time. In the case of this curriculum change, the skill of research design was able to be enhanced and retained significantly through the use of primary literature. Faculty must remain committed to improving introductory biology for majors and be open to implementing and testing new methods

that may enhance scientific literacy skills in order to produce scientifically literate graduates.

Appendix

TOSLS question numbers (from the original instrument) used in this study and how they aligned with each skill category.

TOSLS Skill	TOSLS Question Number
Identify a valid scientific argument	1
Understand elements of research design	4, 25
Make a graph	5
Read and interpret graphical representations of data	2, 6, 7, 18
Solve problems using quantitative skills	16, 20, 23
Understand and interpret basic statistics	3, 19, 24
Justify conclusions based on quantitative data	21, 28

CONCLUSION

Calls for undergraduate biology education reform have been answered at the University of Tennessee by implementing changes to the majors' introductory biology sequence aligned with *Vision and Change* recommendations (AAAS, 2011), including using more active learning within these courses. Changes to concepts, competencies and active learning practices were paired with changes to course structure that included a reduction in lab time and increase in small group discussion time. Instructional practices, faculty perception of active learning use, and student scientific literacy were assessed throughout the reform process.

Chapter one investigated instructional practices during the first year of reform. The instructors of the introductory biology sequence all used at least some active learning pedagogies in their classrooms, meaning that their students likely spent at least some time explicitly thinking about course concepts during class time. However, not all instructors implemented active learning using a collaborative or group context, and there were significant differences between the introductory courses in use of active learning. The baseline data reported in this study were used as feedback to the faculty allowing them to see where they stood in comparison to their own perceptions, others teaching the course, and the sequence as a whole. The data from this study informed future professional development by revealing the need to integrate more group active learning in lecture courses to increase the probability of learning gains for students. The following year, professional development concentrated on how group active learning could be implemented in introductory lecture courses.

Chapter two followed faculty throughout the reform process by observing their instructional practices in the classroom over three years and interviewing them each year about their perceptions of their classroom practices. Unlike the first year of reform, the second year of reform revealed that all instructors were now using group active learning strategies, some for the first time. In lecture alone, instructor active learning use, but not group active learning use, significantly increased over the three years of curriculum reform, although changes between years were not significantly different. The inclusion of discussions in year three significantly increased gains in active learning, group active learning, verbal question, and activity use over time. The interviews provided evidence for how instructional change among the faculty was achieved. There were shifts over the course of the reform in how faculty defined, planned for, and implemented active learning in their classrooms. This may suggest that meaningfully reforming instruction can be accomplished by making small programmatic changes that accumulate over time, thus reducing barriers such as time to plan for change (Dancy & Henderson, 2010). This chapter provided evidence that changes in beliefs can shape practices over time when instructors are engaged in a curriculum reform and suggests that for this to happen, instructors must (1) engage in reflection on what constitutes best practices and their own use of those practices, (2) their reflection must feature a collaborative component, and (3) they must be given time.

Chapter three found that students experiencing the inclusion of discussions featuring acientific literature had a significantly higher gain in scientific literacy compared to the students that did not experience the discussions. While both groups of students gained in scientific literacy after the introductory sequence, higher gains were

achieved by students who experienced the curriculum designed to align with the *Vision and Change* competencies (AAAS, 2011). These students also retained their gains in research design at a higher level as measured by the retention assessment. Assessing the effectiveness of curricula is necessary to ensure that students are achieving the intended outcomes of a course. These data provided information to the program about the acquisition of scientific literacy by introductory students and identified adjustments that could be made to the curriculum.

As a result of the curriculum reform initiatives undertaken at the University of Tennessee, the introductory instructors increased the use of active learning in their classes over time and changed their perceptions of how students learn. Students who experienced the new discussion curricula had greater gains in scientific literacy, and retained more information about research design, despite a loss of lab time. A reform specifically designed to target instructor practices and process of science skills succeeded. Collectively, these studies provide evidence for the importance of collaboration, reflection, and time for instructors to achieve the pedagogical changes needed for modern undergraduate biology reform.

As the fields of psychology and education continue to grow with breakthroughs in how people learn, best classroom practices will continue to evolve. John Dewey once said, “If we teach today’s students as we taught yesterday’s, we rob them of tomorrow.” Instructors need time to continually reflect on and adjust their teaching styles in order to provide students the best education possible. Perhaps the very first step for improving instruction at any education level is for teachers to take the time to give their own instruction some thought.

REFERENCES

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn science. *Science,* 333, 1096 – 1097.

American Association for the Advancement of Science. (1989). *Science for all Americans: Education for a changing future.* Washington, D.C.

American Association for the Advancement of Science. (2010). *Vision and change in undergraduate biology education: A call to action.* Washington, D.C.

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action.* Final report. Washington, D.C.

Andrews, T.M., Leonard, M.J., Colgrove, C.A., & Kalinowski, S.T. (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE-Life Sciences Education, 20*(4), 394-405.

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performances in introductory biology. *CBE-Life Science Education 8,* 203 - 213.

Auerbach, A.J., & Schussler, E.E. (2016). Instructor use of group active learning in an introductory biology sequence. *Journal of College Science Teaching,* 45(*5*), 67 – 74.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How People Learn: Brain, Mind, Experience, and School. Committee on Developments in the Science of Learning*. Washington, DC: National Academies Press.

Buck, L.B., Bretz, S.L., & Towns, M.H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, 5, 52 - 58

Chi, M.T.H., DeLeeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Coleman, E.B., Brown, A.L., & Rivkin, I.D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of Learning Sciences, 6*, 347-365.

Collins, J.W. 3rd, & O'Brien, N.P. (editors). (2003). *The Greenwood Dictionary of Education*. Westport, CT: Greenwood.

Couch, B.A., Brown, T.L., Schelpat, T.J., Graham, M.J., & Knight, J.K. (2015). Scientific teaching: Defining a taxonomy of observable practices. *CBE-Life Sciences Education*, 14, 1-12.

Crossgrove, K., & Curran, K.L. (2008). Using clickers in nonmajors- and majors-level biology courses: Student opinion, learning, and long-term retention of course material. *CBE-Life Sciences Education, 7*(1), 146 – 154.

Dancy, M., & Henderson, C. (2010). Pedagogical practices and instructional change of physics faculty. *American Journal of Physics*, 78, 1056-1063.

Derting, T.L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology: Positive relationships with long-term student achievement. *CBE – Life Sciences Education, 9*, 462-472.

Dirks C., & Cunningham M. (2006). Enhancing diversity in science: is teaching scientific process skills the answer? *CBE-Life Sciences Education*, 5, 218–226.

Dougherty, R.C., Bowen, C.W., Berger, T., Rees, W., Mellon, E.K., and Pulliam, E.

(1995). Cooperative learning and enhanced communication: Effects on student performance, retention, and attitudes in general chemistry. *Journal of Chemical Education, 72*(9), 793 - 797.

Ebert-May, D., Derting, T.L., Hodder, J., Momsen, J.L., Long, T.M., & Jardeleza, S.E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience*, 61, 550-558.

Eddy, S.L., Converse, M., & Wenderoth, M.P. (2015). PORTAAL: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE-Life Sciences Education*, 14, 1-6.

Firooznia, F., & Andreadis, D.K. (2007). Information literacy in introductory biology. *Journal of College Science Teaching,* 35, 23 – 27.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Oloroafor, N., Jordt, H., & Wenderoth, M.P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (*23*), 8410-8415.

Freeman, S., Haak, D., & Wenderoth, M.P. (2011). Increased course structure improves performance in introductory biology. *CBE – Life Sciences Education, 10*, 175-186.

Freeman, S., O'Conner, E.O., Parks, J.W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., & Wenderoth, M.P. (2007). Prescribed active learning increases performance in introductory biology. *CBE-Life Sciences Education,* 6, 132-139.

Glynn, S.M., & Muth, K.D. (1994). Reading and writing to learn science: Achieving

scientific literacy. *Journal of Research in Science Teach*ing, 31: 1057–1073.

Gormally, C., Evans, M., & Brickman, P. (2014). Feedback about teaching in higher ed: Neglected opportunities to promote change. *CBE-Life Sciences Education*, 13, 187-199.

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a Test of Scientific Literacy Skills (TOSLS): measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sciences Education, 11*(4), 364 – 377.

Gottesman, A.J., & Hoskins, S.G. (2013). CREATE cornerstone: Introduction to scientific thinking, a new course for stem-interested freshmen, demystifies scientific thinking through analysis of scientific literature. *CBE-Life Science Education*, 12, 159 – 172.

Haak, D., HilleRislambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*, 1213-1216.

Hake, R.R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66,* 64-74.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., (Miller) Lauffer, S., Stewart, J., Tilghman, S.M., & Wood, W.B. (2004). Scientific teaching. *Science, 304*, 521-522.

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching, 48*(8) 952 – 984.

Henderson, C., & Darcy, M.H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics – Physics Education Research, 3*(2), 020102.

Herman, C.S., Munk, M.H.J., & Engel, A.K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Science, 8*(8), 347-355.

IBM Corporation. (2011). IBM statistics for windows, version 22.0. Armonk, NY: IBM Corp.

Jerde, C.L., & Taper, M.L. (2004). Preparing undergraduates for professional writing. *Journal of College Science Teaching, 33*(7), 34 – 37.

Johnson, D.W., & Johnson, R. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Company.

Khishfe, R. (2015). A look into students' retention of acquired nature of science understandings. *International Journal of Science Education, 37*(10), 1639 – 1667.

Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Knight J.K., & Wood, W.B. (2005). Teaching more by lecturing less. *Cell Biology Education, 4*, 298-310.

Kozeracki C.A., Carey M.F., Colicelli J., & Levis-Fitzgerald M. (2006). An intensive

primary-literature-based teaching program directly benefits undergraduate science majors and facilitates their transition to doctoral programs. *CBE-Life Sciences Education*, 5, 340 – 347.

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Thousand Oaks, CA: Sage.

Lemke, J.L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching, 38*, 296-316.

Michael, J. (2006). Where's the evidence that active learning works? *Adv Physiol Educ* 30: 159-167.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academies Press.

National Research Council. (2004). *Bio2010: Transforming undergraduate education for future research biologists.* Washington, DC: National Academies Press.

National Science Board. (1986). *Undergraduate science, mathematics, and engineering education.* National Science Foundation, Washington, DC.

National Science Foundation. (1996). *Shaping the future: New experiences for undergraduate education in science, mathematics, engineering, and technology. Report of the advisory committee to the NSF directorate for education and human resources*. Washington, DC: National Science Foundation.

O'Day, D.H. (2007). The value of animations in biology teaching: A study of short-term and long-term memory retention. *CBE-Life Sciences Education*, 6, 217-223.

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research,* 62(3), 307 – 332.

PCAST (President's Council of Advisors on Science and Technology). 2012. *Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Washington, DC: Executive Office of the President. Viewed 8 Feb 2012.

Porter, J.A., Wolbach, K.C., Purzycki, C.B., Bowman, L.A., Agbada, E., & Mostrom, A.M. (2010). Integration of Information and Scientific Literacy: Promoting Literacy in Undergraduates. *CBE-Life Sciences Education*, *9*(4), 536–542.

Resnick, L.B., & Resnick, D.P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford & M.C. O'Connor *Changing assessments: Alternative views of aptitude, achievement, and interaction.* (37 – 75). New York: Springer.

Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers.* Los Angeles, CA: SAGE Publications.

Seymour, E. and N.M. Hewitt. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences*. Westview Press: Boulder, CO.

Shaw, G.L., McGaugh, J.L., & Rose, S.P.R. (Eds.). (1990). *Neurobiology of learning and memory.* Teaneck, NJ: World Scientific Publishing.

Silverthorn, D.U., Thorn, P.M., & Svinicki, M.D. (2006). It's difficult to teach the way we teach: lessons from integrative themes in physiology curriculum module project. *Adv Physiol Education, 30*(4), 204 – 214.

Smith, M.K., Jones, F.H.M., Gilbert, S., & Wieman, C.E. (2013). The

classroom observation protocol for undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE-Life Sciences Education,* 12(4), 618 – 627.

Smith, M.K., Vinson, E.L., Smith, J.A., Lewin, J.D., & Stetzer, M.R. (2014). A campus-wide study of STEM courses: New perspectives on teaching practices and perceptions. *CBE-Life Sciences Education, 13*(4), 624 – 635.

Smith, M.K., Wood, W.B., Adams, W.K., Wieman, C., Knight, J.K., Guild, N., & Su, T.T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, 323, 122-124.

Smith, M.K., Wood, W.B., Krauter, K., & Knight, J.K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE-Life Sciences Education*, 10, 55-63.

Spencer, S.J., Steele, C.M., & Quinn, D.M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4 – 28.

Steele, C.M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797 - 811.

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.

Sunal, D.W., Hodges, J., Sunal, C.S., Whitaker, K.W., Freeman, M., Edwards, L., Johnston, R.A., & Odell, M. (2001). Teaching science in higher education: Faculty professional development and barriers to change. *School Science and Mathematics,* 101(*5*), 246 – 257.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using Multivariate Statistics*, 4th ed. Needham Heights, MA: Ally & Bacon.

Turpen, C., & Finkelstein, N.D. (2009). Not all interactive engagement is the same: variations in physics' professors implementation of peer instruction. Physical Review Special Topics – *Physics Education Research*, 5, 1-18.

Turpen, C., & Finkelstein, N.D. (2010). The construction of different classroom norms during peer instruction: students perceive differences. Physical Review Special Topics – *Physics Education Research*, 6, 020123, 1-22.

Webb, N.M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21-39.

Weimer, M. (2002). *Learner-Centered Teaching: Five Key Changes to Practice*, San Francisco, CA: Jossey-Bass.

Welsh, A.J. (2012). Exploring undergraduates' perceptions of the use of active learning techniques in science lectures. *Journal of College Science Teaching, 42,* 80-87.

Westfall, P.H., & Henning, K.S.S. (2013). *Understanding Advanced Statistical Methods.* Boca Raton, FL: CRC Press.

Wiggins, G., & McTighe, J. (1998). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

VITA

Anna Jo Auerbach was born in Decatur, Georgia to the parents of Howard K. Jones and Susan S. Davis. She completed elementary through high school in Douglasville, Georgia. She attended Valdosta State University from 2002 – 2006 graduating in December 2006 with a Bachelor degree in Psychology with a minor in Creative Writing. From there she pursued a Master of Education degree with a focus on School Counseling, graduating in 2009. In 2010, she began her Master of Science in Applied Educational Psychology. She graduated in December, 2012 and that same semester became a student in Ecology and Evolutionary Biology at the University of Tennessee with a research focus on undergraduate biology education. She graduated with a Doctor of Philosophy in Ecology and evolutionary Biology in May, 2016. She is continuing her research experience in a post-doctoral position at the University of Georgia.