

Curriculum Alignment with *Vision and Change* Improves Student Scientific Literacy

Anna Jo Auerbach* and Elisabeth E. Schussler

Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996

ABSTRACT

The *Vision and Change in Undergraduate Biology Education* final report challenged institutions to reform their biology courses to focus on process skills and student active learning, among other recommendations. A large southeastern university implemented curricular changes to its majors' introductory biology sequence in alignment with these recommendations. Discussion sections focused on developing student process skills were added to both lectures and a lab, and one semester of lab was removed. This curriculum was implemented using active-learning techniques paired with student collaboration. This study determined whether these changes resulted in a higher gain of student scientific literacy by conducting pre/posttesting of scientific literacy for two cohorts: students experiencing the unreformed curriculum and students experiencing the reformed curriculum. Retention of student scientific literacy for each cohort was also assessed 4 months later. At the end of the academic year, scientific literacy gains were significantly higher for students in the reformed curriculum ($p = 0.005$), with those students having double the scientific literacy gains of the cohort in the unreformed curriculum. Retention of scientific literacy did not differ between the cohorts.

INTRODUCTION

Scientific literacy can be described as the knowledge and understanding of scientific concepts and processes (American Association for the Advancement of Science [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 technological 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 (National Research Council, 2004; AAAS, 2011), studies that have reported how these curricula impact student proficiencies in scientific literacy are rare. 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 scientific literacy in students (Dirks and Cunningham, 2006; Porter *et al.*, 2010; Gormally *et al.*, 2012), 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.

Peggy Brickman, *Monitoring Editor*

Submitted May 4, 2016; Revised March 3, 2017;

Accepted March 14, 2017

CBE Life Sci Educ June 1, 2017 16:ar29

DOI:10.1187/cbe.16-04-0160

*Address correspondence to: Anna Jo Auerbach (ajone139@vols.utk.edu).

© 2017 A. J. Auerbach and E. E. Schussler.

CBE—Life Sciences Education © 2017 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Non-commercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

Retention studies are difficult, however, because of the need to retest 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), 4 months later (Crossgrove and Curran, 2008), and as long as 3 years later (Derting and Ebert-May, 2010).

Assessments of interventions designed to improve different aspects of student scientific literacy have generally documented short-term improvements in student understanding. Only one study assessing longer-term retention of a scientific literacy aspect was found, and it was conducted with high school students. Kozeracki *et al.* (2006) found that a primary literature-based teaching program that included mentored research along with a seminar course on the presentation and critical analysis of scientific articles benefited student scientific literacy. Porter *et al.* (2010) identified gains in information literacy related to students' ability to access, retrieve, analyze, and evaluate primary scientific literature after participating in an integrated information literacy and scientific literacy exercise in their first-year biology course. Gormally *et al.* (2012) showed that students experiencing a project-based nonmajors' biology course had higher gains in scientific literacy compared with a traditional nonmajors' course. However, retention of literacy skills was not investigated in any of these studies summarized. When retention of skills or knowledge in students is examined, students do not always retain their initial gains. Khishfe (2015), for example, investigated retention of high school students' understanding of the nature of science and found that, 4 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 a large southeastern university undergoing curricular reform to its 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 curricular 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). Before the reform, the courses were each taken for 4 credit hours, had large lectures ($N = 225\text{--}250$) that met for 150 minutes each week, and had an associated 3-hour weekly guided-inquiry lab (Buck *et al.*, 2008) taught by graduate teaching assistants (GTAs). There were no explicit curricula to focus on scientific literacy in the lecture courses, and labs were focused on scientific experimentation skills.

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 new science process skills discussion taken once during the introductory sequence). Students only had to take this course once and could take it either semester. The labs remained in the same format (guided inquiry) and emphasized the same experimentation skills, using half of the same exercises featured in the previous two-semester curricula. At the same time, discussions to more fully explore process skills in each lab exercise were added to the course. Discussions included practicing data collection and data analysis as a group in lab, learning how to use primary literature to support hypotheses,

exploring appropriate scientific communication, and experimental design.

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). Each discussion used scientific articles as the basis for the activities and explorations of the process of science, having students interpret the figures, identify the hypothesis, synthesize the results, and break down the scientific argument being made. A more in-depth description of the three discussion sections, including sample lessons plan and information about how to obtain the full curricula, is provided as Supplemental Material A.

Community members who represented constituents of the course created the discussion curricula using backward design (Wiggins and McTighe, 1998). Members of the community included faculty, postdocs, and 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 the specific process of science learning objectives and using primary literature (Supplemental Material A). Many researchers call for the inclusion of information literacy skills to enable students to achieve scientific literacy skills (Glynn and Muth, 1994; Firooznia and Andreadis, 2007), and others cite using primary literature as a way to build scientific thinking (Gottesman and Hoskins, 2013). The curricula used a scaffolding process in which skills were progressively learned by students and then built upon throughout the courses (Supplemental Material A).

The 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, in which 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 course concepts than traditional lecture, but these studies require a time delay in the measurement of student outcomes (Dougherty *et al.*, 1995; Derting and Ebert-May, 2010).

Assessing the gain and retention of scientific literacy skills as a result of curricular changes is important in order to evaluate the effectiveness of this *Vision and Change* reform to foster new competencies in biology students. This study compared the gains in and retention of scientific literacy skills for students in the 2013–2014 introductory sequence who experienced the two-course sequence with two semesters of lab and did not experience scientific literacy discussions (cohort 1) with those of students from the 2014–2015 cohort who experienced the two-course sequence with one semester of an independent lab course and the three associated scientific literacy discussions (cohort 2). This study addressed the following questions:

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

METHODS

Participants

The participants were undergraduate students at a large southeastern university registered for all courses of either the majors' introductory biology sequence in the 2013–2014 or 2014–2015 academic years. Both cohorts had approximately 650 students starting in Fall 2013 and 2014 as potential participants. A total of 305 participants were included in this study. The first cohort (N_1) had 156 participants and the second cohort (N_2) had 149 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 Review Board approved the research before data collection.

Instrument

The Test of Scientific Literacy Skills (TOSLS; Gormally *et al.*, 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). The original test was designed to be implemented in 35 minutes. Owing to class time constraints (20 minutes were available), only a portion of the TOSLS was used for this study.

The TOSLS items used were selected by letting the faculty associated with the courses choose 16 TOSLS questions that best represented the competencies outlined within their courses. These items aligned with seven of the nine skills on the original instrument but did not necessarily include all of the items for each skill category (Supplemental Material B). The TOSLS skill categories and number of items from each category chosen were: identify a valid scientific argument (one of the three items were selected), understand elements of research design (two of the four items were selected), make a graph (one of one item selected), interpret a graph (four of four items selected), solve problems using quantitative skills (three of three items selected), understand and interpret basic statistics (three of three items selected), and justify conclusions based on quantitative data

(two of two items selected). Overall scientific literacy and scores on each question were measured through the TOSLS implementation. It should be noted that the way we selected and used items for this study was different from the original, validated TOSLS instrument.

All administrations of the TOSLS used the same 16 questions that were selected by the faculty group. Student scores on the TOSLS assessment overall were calculated by summing the number of correct survey item responses; thus student scores ranged from 0 to 16 (0% to 100%). Cohort scores for each question were also calculated based on the number of correct responses for that particular question.

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 1). Cohort 1 completed this assessment schedule in 2013 and 2014, and cohort 2 completed it in 2014 and 2015. The first cohort took the pretest during their first lab class period in the first introductory course. The second cohort took theirs during their first discussion class period of the first introductory course. The postassessment was given to cohort 1 during the last lab class period of the second introductory course, and cohort 2 took the postassessment during their last discussion class period of their second introductory course. Each time, the TOSLS was administered at the beginning of class. The GTAs teaching the labs or discussions administered the pre- and postassessments by using a scripted introduction and set of instructions for implementation. Students filled out a demographic sheet after the completion of the TOSLS to prevent stereotype threat (Steele and Aronson, 1995; Spencer *et al.*, 1999). The survey instructions were presented by A.J.A. 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 midlength retention testing time of 4 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 A.J.A. during the first lab and discussion periods of each of these three classes in Fall 2014. Completed assessments were sorted to identify students who had taken both the pre- and postassessments as part of cohort 1. This method produced a very low number of matches (20 out of the 156 matched pre/postassessments for cohort 1), 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

TABLE 1. Schedule of assessments of scientific literacy skills via the TOSLS for each cohort

Group ^a	Fall 2013	Spring 2014	Fall 2014	Spring 2014	Fall 2015
Cohort 1	Pre (1)	Post (2)	Retention (3)		
Cohort 2			Pre (1)	Post (2)	Retention (3)

^aCohort 1 experienced two courses with associated labs and no discussion sections over their introductory sequence. Cohort 2 experienced one stand-alone lab course with a discussion and two course-associated discussion sections over their introductory sequence.

the retention sample for the second cohort, we used a different recruitment approach in Fall 2015. All students who had completed the pre- and postassessments for cohort 2 ($N_2 = 149$) were invited by email to participate in an online 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 online 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 postassessments were matched using student identification numbers. All nonmatching data (e.g., a student who took the pre- but not the postassessment) 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. Eight participants were eliminated from the second cohort's retention sample because they completed the survey in less than 5 minutes, which was judged to be insufficient time to adequately read and complete the assessment. For each administration of the TOSLS, a KR-20 was calculated as a measure of internal consistency reliability. The data were checked to ensure that all values were within the specified range of 0–16, assessed for outliers, and assessed for normality.

The first research question was to identify whether students experiencing the new curriculum gained in scientific literacy as compared with students who experienced the prior curriculum. Scores were calculated as a percentage of the 16 questions that were answered correctly by cohort 1 and cohort 2 on the pre- and postassessments. Using these data, independent-samples t tests were conducted comparing normalized gain from pre–post scores for each cohort and also to compare the two cohorts' gain by individual TOSLS questions. Comparing normalized gain is a commonly used method in active-learning research (Hake, 1998; Knight and Wood, 2005). All analyses were conducted using SPSS (IBM Corporation, 2011).

Student scores were also compared using student demographics to ensure equity among students within cohorts. Owing to low numbers within the category of ethnicity, the variables were reduced to white and nonwhite students (Tabachnick and Fidell, 2001). Two separate mixed analyses of variance (ANOVAs) with gender (male/female) and ethnicity (white/nonwhite) as the independent variables and pre score and post score as the dependent variables were calculated. A repeated-measures ANOVA was used to determine whether there were differences in pre scores and post scores for students by year in school at the university.

The second research question was to identify whether students experiencing the new curriculum had improved retention of scientific literacy as compared with 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 scores for students who participated in the pre- and postassessment and the retention

measure. The same methods were used to calculate student scores on retention of scientific literacy overall. Using these data, independent-samples t tests were conducted comparing the overall change in post retention scores for each cohort and also to compare the two cohorts' gain by individual TOSLS questions. All analyses were conducted using SPSS (IBM Corporation, 2011).

Two separate repeated-measures ANOVAs with gender (male/female) and ethnicity (white/nonwhite) as the independent variables with post score and retention as the dependent variables were calculated. A repeated-measures ANOVA was used to determine whether there were differences in post score and retention for students by year in school at university. Effect sizes were calculated using Cohen's d for a repeated-measures design due to the correlation of pre to post scores and post to retention scores (Dunlap *et al.*, 1996).

RESULTS

Both cohorts revealed a significant gain in scientific literacy overall from pre- to posttest. The first cohort showed a gain of 7% in scientific literacy scores from pretest ($M = 0.64$, $SD = 0.17$) to posttest ($M = 0.71$, $SD = 0.17$; Table 2), $t(19) = 2.14$, $p = 0.046$; while the second cohort showed a gain of 13% from pretest ($M = 0.58$, $SD = 0.17$) to posttest ($M = 0.71$, $SD = 0.71$), $t(49) = 2.419$, $p = 0.000$. Both cohorts' retention scores ($M_1 = 0.71$, $SD_1 = 0.17$; $M_2 = 0.70$, $SD_2 = 0.20$) were similar to their post scores ($M_1 = 0.76$, $SD_1 = 0.13$; $M_2 = 0.71$, $SD_2 = 0.15$).

Not all students who took the pretest ($N_1 = 610$; $N_2 = 602$) also took the posttest ($N_1 = 490$; $N_2 = 394$). This was due to student absences on the first day of either course, students who took the sequence out of order (which was discouraged but allowable within the curriculum), or did not complete the sequence in the same academic year. However, both cohorts had similar numbers of students with matched pre- and postassessments ($N_1 = 156$; $N_2 = 149$; Table 3). 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 average for 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 freshmen. There were also sophomores (24.7%), juniors (11.1%), and seniors (3.0%) in the sample. KR-20 analysis indicated the shortened instrument reliability ranged from 0.54 to 0.76, with the lower reliability being for the retention measures with the small student sample size (values for each implementation shown on Tables 2 and 4).

Scientific Literacy Gain

When comparing the cohorts' gains in scientific literacy overall from pretest to posttest, a t test revealed that cohort 2 experienced a significantly higher gain than cohort 1, $t(303) = -2.837$, $p = 0.005$ (Figure 1), allowing cohort 2 to attain similar scientific literacy by the posttest, despite having significantly lower pre scores. 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.

TABLE 2. Average (M) TOSLS pre score and post score reported with SD for overall scientific literacy (SL) and by individual question^a

Overall SL (16 items)	Cohort ^b (N)	Pre score M(SD)	Post score M(SD)	Gain	Effect size ^c (d)
Overall SL*	1 (156)	0.64 (0.17)	0.71 (0.17)	0.07	0.39
	2 (149)	0.58 (0.17)	0.71 (0.17)	0.13	0.71
Skill related to individual question (TOSLS question no.)					
Identify a valid scientific argument (1)*	1 (156)	0.83 (0.38)	0.91 (0.29)	0.08	0.25
	2 (149)	0.87 (0.34)	0.84 (0.37)	-0.03	-0.07
Read and interpret graphical representations of data (2)	1 (156)	0.57 (0.5)	0.66 (0.48)	0.09	0.18
	2 (149)	0.53 (0.50)	0.61 (0.48)	0.08	0.16
Read and interpret graphical representations of data (6)	1 (156)	0.84 (0.37)	0.82 (0.39)	-0.02	-0.05
	2 (149)	0.82 (0.39)	0.87 (0.33)	0.05	0.15
Read and interpret graphical representations of data (7)*	1 (156)	0.85 (0.36)	0.78 (0.41)	-0.07	-0.17
	2 (149)	0.77 (0.42)	0.83 (0.37)	0.06	0.15
Read and interpret graphical representations of data (18)	1 (156)	0.72 (0.45)	0.63 (0.49)	-0.1	-0.21
	2 (149)	0.67 (0.47)	0.68 (0.46)	0.01	0.01
Understand and interpret basic statistics (3)*	1 (156)	0.52 (0.50)	0.83 (0.37)	0.31	0.7
	2 (149)	0.59 (0.49)	0.78 (0.42)	0.19	0.4
Understand and interpret basic statistics (19)	1 (156)	0.62 (0.49)	0.81 (0.40)	0.19	0.46
	2 (149)	0.59 (0.49)	0.76 (0.43)	0.17	0.35
Understand and interpret basic statistics (24)	1 (156)	0.42 (0.50)	0.56 (0.50)	0.14	0.28
	2 (149)	0.50 (0.50)	0.54 (0.50)	0.04	0.08
Understand elements of research design (4)	1 (156)	0.66 (0.48)	0.74 (0.44)	0.08	0.18
	2 (149)	0.62 (0.49)	0.77 (0.42)	0.15	0.33
Understand elements of research design (25)*	1 (156)	0.77 (0.42)	0.74 (0.44)	-0.05	-0.09
	2 (149)	0.13 (0.34)	0.77 (0.42)	0.64	1.76
Make a graph (15)	1 (156)	0.46 (0.50)	0.60 (0.50)	0.14	0.29
	2 (149)	0.50 (0.50)	0.70 (0.46)	0.2	0.42
Solve problems using quantitative skills (16)	1 (156)	0.74 (0.44)	0.77 (0.42)	0.03	0.06
	2 (149)	0.64 (0.48)	0.72 (0.45)	0.08	0.17
Solve problems using quantitative skills (20)	1 (156)	0.38 (0.49)	0.55 (0.50)	0.16	0.34
	2 (149)	0.34 (0.47)	0.45 (0.50)	0.11	0.22
Solve problems using quantitative skills (23)	1 (156)	0.75 (0.43)	0.72 (0.45)	-0.03	-0.07
	2 (149)	0.74 (0.44)	0.75 (0.43)	-0.01	0.01
Justify inferences, predictions, and conclusions based on quantitative data (21)	1 (156)	0.60 (0.49)	0.69 (0.46)	0.09	0.19
	2 (149)	0.58 (0.49)	0.75 (0.44)	0.17	0.37
Justify inferences, predictions, and conclusions based on quantitative data (28)	1 (156)	0.49 (0.50)	0.55 (0.50)	0.06	0.11
	2 (149)	0.46 (0.50)	0.46 (0.50)	0	0

^aIndividual questions are labeled by the skill in which they are related; similar skills are grouped together.

^bCohort 1 ($N = 156$) experienced two courses with associated labs and no discussion sections over their introductory sequence. Cohort 2 ($N = 149$) experienced one stand-alone lab course with a discussion and two course-associated discussion sections over their introductory sequence.

^cEffect sizes are reported as Cohen's d . The KR-20 for the cohort 1 pretest was 0.66 and was 0.71 for the posttest. The KR-20 for the cohort 2 pretest was 0.74 and was 0.76 for the posttest.

* $p < 0.05$ (indicates a significant difference between cohorts when comparing gains).

Separate t tests were used to determine whether the cohorts differed in pre- to posttest gains in scores on individual questions. These analyses revealed significant differences between cohorts on four questions, one question (1) showed a gain for cohort 1 and a loss for cohort 2, $t(303) = 2.282$, $p = 0.023$; two questions (7, 25) revealed a loss for cohort 1 and gain for cohort 2, $t(303) = -1.975$, $p = 0.049$ and $t(303) = -10.864$, $p = 0.000$ respectively; and the remaining question (3) showed gains for both cohorts with cohort 1 achieving a

higher gain compared with cohort 2, $t(303) = 2.017$, $p = 0.045$ (Table 2).

Scientific Literacy Retention

Four months after the posttest, both the first and second cohort had slightly lower scientific literacy scores on the overall retention test. The loss was slightly larger for the first cohort ($d = -0.41$; Table 4) compared with the second cohort ($d = -0.01$; Table 4), but these differences were not significant

TABLE 3. Percentages of student group demographics of gender, ethnicity, and year in school by cohort^a

	Cohort 1		Cohort 2	
	Pre-Post (N = 156)	Retention (N = 20)	Pre-Post (N = 149)	Retention (N = 50)
Gender				
Female	56%	55%	58%	72%
Male	44%	40%	42%	26%
No response	1%	5%	1%	2%
Ethnicity				
White	78%	60%	80%	86%
Nonwhite	21%	40%	18%	12%
No response	1%	—	2%	2%
Year in school				
Freshman	59%	60%	60%	68%
Sophomore	26%	25%	24%	24%
Junior	10%	10%	13%	8%
Senior	5%	5%	3%	—

^aCohort 1 experienced two courses with associated labs and no discussion sections over their introductory sequence. Cohort 2 experienced one stand-alone lab course with a discussion and two course-associated discussion sections over their introductory sequence. Due to low numbers within the category of ethnicity, the variables were reduced to white and nonwhite students to allow for analysis. Demographics marked with a dash (—) indicate no participants.

(Figure 2). There were also no differences in retention of scientific literacy for students when comparing gender and ethnicity or between students in different years.

Separate *t* tests were used to determine whether the cohorts differed in scores from posttest to retention by individual question. Significant differences were found for two questions (3, 25). For question 3, cohort 1 gained from posttest to retention, while cohort 2 experienced a loss, $t(68) = 2.051, p = 0.044$. For question 25, cohort 1 experienced a large loss, while cohort 2 only experienced a slight loss, $t(68) = -4.605, p = 0.000$.

DISCUSSION

Students in both cohorts experienced significant gains in scientific literacy overall over one academic year as measured by TOSLS. Comparison of the cohorts revealed greater overall gains by the students who experienced the curriculum aligned with the *Vision and Change* competencies (AAAS, 2011), allowing their post scores to be similar to those of the first cohort, despite starting with lower scientific literacy. Students in both cohorts were equally able to retain most of their overall scientific literacy gains 4 months after the posttest. When student scores on individual items related to particular scientific literacy skills were examined, the second cohort gained significantly more over one academic year than the first cohort on two items, one related to the skill of understanding elements of research design and the other related to the skill of reading and interpreting graphical representations of data. Cohort 2 also retained significantly higher scores on the latter item compared with the first cohort when assessed 4 months later. The first cohort experienced a higher gain than the second cohort on one question assessing the skill of identifying a valid scientific argument over one academic year, but there were no differences among the cohorts on this question 4 months later.

The first-semester lecture discussion had a focus on diagramming the research designs of the primary literature papers that students read as part of their activities (Supplemental Material A). These diagrams were then explicitly aligned with the hypothesis and result figures shown in the scientific paper

being read. If the performance on the research design TOSLS item is an indicator of better performance on this skill generally, then it may be that this visualization of the research design is an important step in fostering student understanding of this critical skill (Ainsworth *et al.*, 2011). The sample from cohort 2 demonstrated a greater retention of performance on this question 4 months after the course (and for most students there was an 8-month gap from the first-semester discussion to the retention test date), suggesting that the practices may have been effective in embedding this understanding in students' minds. It is interesting to note that a skill such as research design would be expected to arise from laboratory study, and yet cohort 2 had one semester less of laboratory study than cohort 1. It may be that the study of primary literature in conjunction with the design of laboratory experiments could be a more effective mechanism for developing this skill than lab experiences alone.

The ability to identify a valid scientific argument was chosen by faculty as an important part of student scientific literacy, yet the discussion cohort (cohort 2) was significantly different from the students in the nonreformed curriculum in their performance on this question. The nonreformed curriculum students displayed a gain on this item, while the discussion cohort displayed a decrease in that measure of their ability. This was particularly surprising because the second-semester lecture discussion focused almost exclusively on scientific argumentation as a central learning objective. It should also be noted that the differences between cohorts on this question disappeared by the retention time point, indicating an equal performance 4 months later from those samples of students.

The second cohort displayed nearly twice the gain in scientific literacy skills over the academic year when compared with the first cohort; however, both cohorts were equally able to retain their respective overall gains in scientific literacy as measured 4 months after the completion of the courses. It should be noted that the two cohorts had similar posttest scores, indicating that the difference in gain was due to cohort 2 beginning lower than cohort 1 in scientific literacy but catching up by the end of the year. It is encouraging to know that

TABLE 4. Average (M) TOSLS post score and retention score reported with SD (SD) for overall scientific literacy (SL) and by individual question^a

Overall SL (16 items)	Cohort ^b	Post score	Retention score	Effect size ^c	
	(N)	M(SD)	M(SD)	Gain	(d)
Overall SL (16 items)	1 (20)	0.76 (0.13)	0.71 (0.17)	-0.05	-0.22
	2 (50)	0.71 (0.15)	0.70 (0.20)	-0.01	-0.04
Skill related to individual question (TOSLS question #)					
Identify a valid scientific argument (1)	1 (20)	0.90 (0.31)	0.95 (0.22)	0.05	0.19
	2 (50)	0.84 (0.37)	0.84 (0.37)	0	0
Read and interpret graphical representations of data (2)	1 (20)	0.80 (0.41)	0.60 (0.50)	-0.2	-0.43
	2 (50)	0.66 (0.48)	0.60 (0.50)	-0.06	-0.12
Read and interpret graphical representations of data (6)	1 (20)	0.85 (0.37)	0.75 (0.44)	-0.1	-0.24
	2 (50)	0.78 (0.42)	0.76 (0.43)	-0.02	-0.05
Read and interpret graphical representations of data (7)	1 (20)	0.85 (0.37)	0.70 (0.47)	-0.15	-0.36
	2 (50)	0.92 (0.27)	0.78 (0.42)	-0.16	-0.4
Read and interpret graphical representations of data (18)	1 (20)	0.75 (0.44)	0.75 (0.09)	0	0
	2 (50)	0.66 (0.48)	0.78 (0.05)	0.12	0.27
Understand and interpret basic statistics (3)*	1 (20)	0.85 (0.37)	0.95 (0.22)	0.1	0.31
	2 (50)	0.74 (0.44)	0.66 (0.48)	-0.08	-0.17
Understand and interpret basic statistics (19)	1 (20)	0.75 (0.44)	0.80 (0.41)	0.05	0.12
	2 (50)	0.65 (0.49)	0.84 (0.37)	0.1	0.24
Understand and interpret basic statistics (24)	1 (20)	0.75 (0.44)	0.50 (0.51)	-0.25	-0.52
	2 (50)	0.56 (0.50)	0.56 (0.50)	0	0
Understand elements of research design (4)	1 (20)	0.90 (0.31)	0.95 (0.22)	0.05	0.19
	2 (50)	0.78 (0.42)	0.68 (0.47)	-0.1	-0.22
Understand elements of research design (25)*	1 (20)	0.80 (0.41)	0.10 (0.31)	-0.7	-1.94
	2 (50)	0.78 (0.42)	0.76 (0.43)	-0.02	-0.05
Make a graph (15)	1 (20)	0.55 (0.51)	0.55 (0.51)	0	0
	2 (50)	0.70 (0.46)	0.60 (0.49)	-0.1	-0.21
Solve problems using quantitative skills (16)	1 (20)	0.80 (0.41)	0.85 (0.37)	-0.05	0.13
	2 (50)	0.82 (0.38)	0.86 (0.35)	-0.16	0.11
Solve problems using quantitative skills (20)	1 (20)	0.65 (0.49)	0.60 (0.50)	-0.05	-0.1
	2 (50)	0.46 (0.50)	0.56 (0.50)	0.1	0.2
Solve problems using quantitative skills (23)	1 (20)	0.70 (0.47)	0.90 (0.31)	0.2	0.51
	2 (50)	0.76 (0.43)	0.82 (0.39)	0.06	0.15
Justify inferences, predictions, and conclusions based on quantitative data (21)	1 (20)	0.80 (0.41)	0.85 (0.37)	0.05	0.13
	2 (50)	0.72 (0.45)	0.72 (0.45)	0	0
Justify inferences, predictions, and conclusions based on quantitative data (28)	1 (20)	0.50 (0.51)	0.50 (0.51)	0	0
	2 (50)	0.50 (0.51)	0.48 (0.50)	-0.02	-0.04

^aIndividual questions are labeled by the skill in which they are related; similar skills are grouped together.

^bCohort 1 (N = 20) experienced two courses with associated labs and no discussion sections over their introductory sequence. Cohort 2 (N = 50) experienced one stand-alone lab course with a discussion and two course-associated discussion sections over their introductory sequence.

^cEffect sizes are reported as Cohen's *d*. The KR-20 for the cohort 1 posttest was 0.59 and for the retention test was 0.54. The KR-20 for the cohort 2 posttest was 0.62 and for the retention test was 0.61.

**p* < 0.05 (indicates a significant difference between cohorts when comparing gains).

the students who were assessed retained a good amount of the scientific literacy that they obtained from their introductory course work. 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 everyday scenarios that require evaluating and using data to make informed decisions (AAAS, 2011); therefore, students' retention of what they learn in school is important.

Previous studies (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 literature base, showing that discussion curricula designed to foster scientific literacy skills in introductory courses for biology

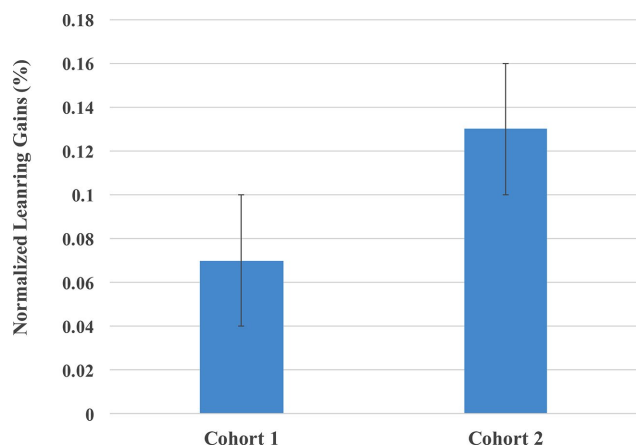


FIGURE 1. Normalized learning gains show the relative gain of scientific literacy from pretest to posttest. Gains are displayed for scientific literacy overall for both cohort 1 (2013–2014; nonreformed curriculum; $N = 156$) and cohort 2 (2014–2015 reformed curriculum; $N = 149$).

majors can increase scientific literacy and that these effects can be retained as long as 4 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 it has shown that these improvements can be achieved through active-learning exercises centered on scientific literature and not necessarily through increasing laboratory experiences. Curricula designed to intentionally enhance a particular skill set can indeed make a difference in outcomes for introductory biology students.

While these results are encouraging to those considering a *Vision and Change* curricular reform, there are some limitations. As mentioned in the *Methods*, the TOSLS instrument could not be used in its entire validated form, which may have impacted this measure of scientific literacy for this study. For each administration of our shortened TOSLS, a KR-20 was calculated, and it was found to range from 0.66 to 0.76 for the pre and post implementations and 0.54 to 0.62 for the retention implementation. The published full instrument reported a reliability of between 0.731 and 0.748 (Gormally *et al.*, 2012). The discrepancy of reliability for TOSLS in this study is likely due to the small number of participants in the retention measures and not using the instrument in its entirety. A score of 0.70 or above is considered acceptable (Cronbach, 1951), because, as this value declines, it is more likely that differences can be attributed to other factors besides student performance on the assessments being given. Therefore, readers should take this potential reliability issue into consideration. However, there is debate in the education community about how much meaning to ascribe to this measure (Dunn *et al.*, 2014). Using student performance on individual question items to infer changes in skill sets is also problematic. Certainly, it is important to remember that performance on particular items on TOSLS should not be taken as a measure of that skill, *per se*, but rather as performance on those particular questions only. In this case, we used comparisons of performance on questions

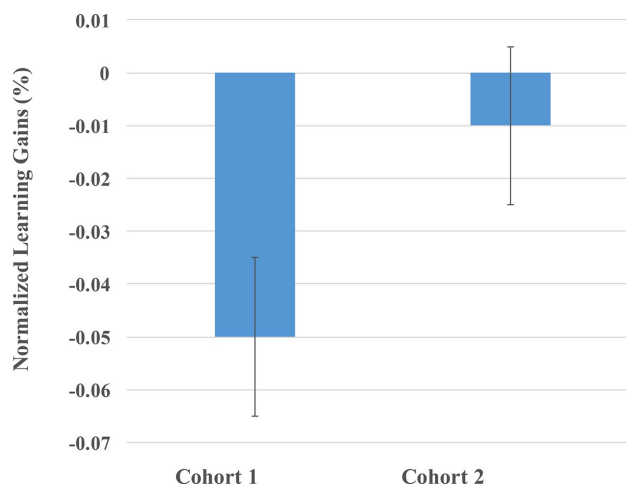


FIGURE 2. Normalized learning gains show the relative retention of scientific literacy from posttest to retention measure and are reported for both cohort 1 (2013–2014; nonreformed curriculum; $N = 20$) and cohort 2 (2014–2015 reformed curriculum; $N = 50$).

between cohorts to make inferences about relative impacts of the curriculum being tested.

There are additional cautionary notes regarding outside influences on student scientific literacy and the difficulty in implementing the retention testing. 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 skills. There was also no way to control for other science courses students were taking over the academic year that may have impacted the results. Regarding the retention measure, the low number of participants within the retention sample make it impossible to generalize these results even to the students within the cohorts. There were also differences in the composition of the retention sample in terms of gender and ethnicity compared with the pre-post sample for both cohorts. This may indicate that the students sampled between the academic year pre/posttests and retention measures were different. There were also differences in the methods used to obtain the retention data, likely resulting in 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 who are given an incentive to take a test online, although it was encouraging that the retention measures were similar despite the different testing contexts. 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.

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 curricular changes such as these is necessary to ensure that students are achieving the intended outcomes of a course or set of courses (Resnick and 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 students will retain some of these skills over time. In the

case of this curricular change, student performances on questions related to research design were enhanced and retained significantly more through the use of primary literature-based discussions versus increasing lab time. 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.

ACKNOWLEDGMENTS

We thank all of the faculty, staff, GTAs, and students for their vital contributions to making this research a reality.

REFERENCES

- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn science. *Science*, *333*, 1096–1097.
- American Association for the Advancement of Science. (AAAS) (1989). *Science for all Americans: Education for a changing future*, Washington, DC.
- AAAS. (2011). *Vision and change in undergraduate biology education: A call to action. Final report*. Washington, DC.
- 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.
- Cronbach, J. L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.
- 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*, 146–154.
- 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? *Cell Biology Education*, *5*, 218–226.
- Dougherty, R. C., Bowen, C. W., Berger, T., Rees, W., Mellon, E. K., & Pulliam, E. (1995). Cooperative learning and enhanced communication: Effects on student performance, retention, and attitudes in general chemistry. *Journal of Chemical Education*, *72*, 793–797.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*, 170–177.
- Dunn, T. J., Baguley, T., & Brunnsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399–412.
- 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., Olorofofor, 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*, 8410–8415.
- Glynn, S. M., & Muth, K. D. (1994). Reading and writing to learn science: Achieving scientific literacy. *Journal of Research Science Teacher*, *31*, 1057–1073.
- 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*, 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 Sciences Education*, *12*, 159–172.
- 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.
- IBM Corporation. (2011). *IBM statistics for Windows. Version 22.0*, Armonk, NY.
- Khishfe, R. (2015). A look into students' retention of acquired nature of science understandings. *International Journal of Science Education*, *37*, 1639–1667.
- 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.
- National Research Council. (2004). *BIO2010: Transforming undergraduate education for future research biologists*, Washington, DC: National Academies Press.
- 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.
- 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*, 536–542.
- Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: new tools for educational reform. In Gifford, B.R., & O'Connor, M.C. (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and interaction* (pp. 37–75). New York: Springer.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Experimental Social Psychology*, *69*, 797–811.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.), Needham Heights, MA: Allyn & Bacon.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.