CBE—Life Sciences Education
Vol. 13, 444–452, Fall 2014

Article

Promoting Inquiry-Based Teaching in Laboratory Courses: Are We Meeting the Grade?

Christopher Beck,*† Amy Butler,† and Karen Burke da Silva†

*Department of Biology, Emory University, Atlanta, GA 30322; †School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia

Submitted December 23, 2013; Revised May 2, 2014; Accepted May 5, 2014
Monitoring Editor: Marshall Sundberg

Over the past decade, repeated calls have been made to incorporate more active teaching and learning in undergraduate biology courses. The emphasis on inquiry-based teaching is especially important in laboratory courses, as these are the courses in which students are applying the process of science. To determine the current state of research on inquiry-based teaching in undergraduate biology laboratory courses, we reviewed the recent published literature on inquiry-based exercises. The majority of studies in our data set were in the subdisciplines of biochemistry, cell biology, developmental biology, genetics, and molecular biology. In addition, most exercises were guided inquiry, rather than open ended or research based. Almost 75% of the studies included assessment data, with two-thirds of these studies including multiple types of assessment data. However, few exercises were assessed in multiple courses or at multiple institutions. Furthermore, assessments were rarely based on published instruments. Although the results of the studies in our data set show a positive effect of inquiry-based teaching in biology laboratory courses on student learning gains, research that uses the same instrument across a range of courses and institutions is needed to determine whether these results can be generalized.

INTRODUCTION

Inquiry-based approaches should be transforming the way science faculty members teach and undergraduate students learn (National Research Council [NRC], 2003a). Since 1985, the National Science Foundation (NSF) has promoted the enhancement of undergraduate science instruction, especially in laboratory courses (Tuss *et al.*, 1998). A push to increase active teaching and learning within undergraduate biology courses occurred in 2003, with the publication of *BIO2010* (NRC, 2003a). In the same year, the NRC called for an increase in the scientific assessment of teaching to improve student learning (NRC, 2003b). A year later, Handelsman and colleagues' seminal paper in *Science* echoed this push for more active learning in science courses, providing examples of scientific teaching in both lecture and laboratory classrooms (Handelsman *et al.*, 2004). They also reiterated the need to rigorously assess our teaching. More recently, the *Vision and Change* report has once again emphasized the need to incorporate inquiry throughout the undergraduate biology curriculum and to use scientific approaches to assess faculty teaching and student learning (American Association for the Advancement of Science [AAAS], 2011). However, a recent review of curricular innovations in undergraduate science courses (both lecture and laboratory) suggests that teaching practices in undergraduate science and engineering courses are changing in the direction of inquiry but, unfortunately, are still not necessarily being assessed rigorously (Ruiz-Primo *et al.*, 2011).

The infusion of inquiry-based approaches into the undergraduate biology curriculum is especially important in laboratory courses or within the laboratory component of a course, because students are applying the process of science in these courses (AAAS, 2011) and are learning to write in an inquiry-based way (Tuss *et al.*, 1998). Surveys that took place more than 20 yr ago suggest that inquiry-based and

DOI: 10.1187/cbe.13-12-0245
Address correspondence to: Christopher Beck (christopher.beck@emory.edu).

open-ended laboratory exercises were used infrequently (10% of courses; Sundberg and Armstrong, 1993). A more recent survey reporting results from 65 colleges and universities in the United States found close to 80% of institutions are using inquiry-based approaches in their laboratory courses (Sundberg *et al.*, 2005), indicating a very promising trend. Yet the Sundberg *et al.* (2005) study did not determine the degree to which these approaches had been implemented or the effectiveness of these inquiry-based teaching methods on student performance. More than 7 yr on, information about inquiry-based approaches is still lacking, with the recent National Research Council report on discipline-based education research rating the evidence in support of inquiry-based learning in laboratory courses in biology as "limited" (NRC, 2012).

Given the repeated calls for increased inquiry-based teaching and learning in the undergraduate biology curriculum and the important role of laboratory courses in the curriculum, we conducted a meta-analysis of peer-reviewed, published, undergraduate inquiry-based laboratory exercises produced since the publication of the *BIO2010* report (NRC, 2003a) and Handelsman *et al.*'s (2004) seminal paper on scientific teaching. Specifically, we addressed two main questions:

1. To what degree have inquiry-based laboratory exercises been developed and published in different subdisciplines in biology, at what level (i.e., nonmajors, introductory, upper-level majors), and what type of inquiry (i.e., guided inquiry, open-ended inquiry, and research)?
2. To what degree have these exercises been assessed and are they in line with the recommendations of the NRC (2003b) and Handelsman *et al.* (2004)?

The first question is important, as it tells us about where future efforts need to be made in curriculum development and publication/dissemination. The second question is important, as it tells us whether laboratory exercises are being assessed, how they are being assessed, and what the assessment tells us about the effectiveness of the exercises. The data suggest where future efforts need to be made in assessment and provide a snapshot of the evidence for the value of inquiry-based laboratory exercises in undergraduate biology courses.

METHODS

We searched on "inquir* lab*" (where * is a wildcard) in the topic field of Web of Science and then limited the results to the years from 2005 to 2012 and to the subject areas of education, educational research, and all areas of biology. The same search was carried out in the Education Resources Information Center (ERIC; www.eric.ed.gov) using ProQuest, which returned some of the same references as those from the Web of Science search. As papers on research in laboratory courses sometimes do not use the term "inquiry," we also searched on "research lab*" in Web of Science and ERIC and limited the results as described above. Finally, we searched "research lab*" in the tables of contents of journals that were found in our previous searches. We combined the results of all searches and then manually excluded all studies that were not explicitly about an inquiry-based laboratory exercise in biology at the undergraduate level. After the exclusion process, 142 studies remained, which were scored for the current study (see the Supplemental Material). Admittedly, our search did not retrieve all publications of inquiry-based laboratory exercises in biology, as these exercises might not be published in journals catalogued by Web of Science or ERIC. However, those studies that include assessment data are perhaps more likely to be published in catalogued journals.

Because we were interested in changes in the publication of inquiry-based laboratory exercises and the assessment of those exercises following the publication of the *BIO2010* report (NRC, 2003a), the NRC's report on assessment of teaching and learning (NRC, 2003b), and Handelsman and colleagues' paper on scientific teaching (Handelsman *et al.*, 2004), we only considered published papers from the United States, as these reports likely had the greatest effect in their country of origin. In setting the time windows for our analysis, we allowed for 2 yr after the reports and 1 yr after Handelsman *et al.*'s paper to allow the community to incorporate the recommendations into their work and for new studies to be published. As a result, we considered studies from 2005 to 2012 as those that might reflect the impact of these publications.

We coded each paper in the following categories:

1. Area of biology: biochemistry, cell biology, developmental biology, genetics, and molecular biology (BCDGMB); ecology and evolutionary biology (EEB); plant and animal anatomy and physiology (organismal biology).
2. Inquiry type: guided inquiry, open-ended inquiry, research (including teacher-collaborative and structured research experiences), and other (e.g., case studies, problem-based learning, model building). We based inquiry types on D'Avanzo (1996) and Weaver *et al*. (2008). In guided inquiry, faculty members provide students with the research question of interest and guide students to an appropriate experimental design. Open-ended inquiry is less structured, with students posing the research question and developing the appropriate methods to address the question. For studies classified as research, the degree of student independence in defining the question and methodology varied. Unlike guided inquiry, the outcomes of the experiments are not known to either the students or faculty. Unlike open-ended inquiry, in which the emphasis is placed on students being in control of the process, students and faculty work collaboratively in research experiences.
3. Course level: nonmajors, introductory majors, upper-level majors, and mixed nonmajors and majors.
4. Assessment type: qualitative (i.e., student comments, interviews), student self-assessment (self-efficacy and student assessment of learning gains), disciplinary content, and other (e.g., scientific reasoning, experimental design, information literacy, statistical literacy).
5. Whether the assessment was based on a published instrument.
6. Whether the paper presented data for a control group.
7. The number of students studied.
8. The results of the study.

For studies in which the laboratory experiments were in more than one area of biology, the study was included for each area of biology covered. Studies that reported multiple types of assessment data were included once for each type of assessment for appropriate analyses. As a result of these multiple codings, the sample sizes for "subdiscipline in biology" and "assessment" are greater than the total number of studies in the data set. In most cases, our coding followed that stated in the paper. In cases in which papers did not provide the required information, we coded them based on the information provided or coded them as "not provided." As a result, the sample sizes vary among analyses.

Data Analysis

All of our data are frequency data (i.e., counts of the number of studies that fall into particular categories). As a result, we used chi-square tests of independence (cross-tabulation) for all analyses. The results of chi-square tests can be biased if the expected count in a particular cell is low. Randomization tests have been proposed as an alternative approach for estimating p values in these cases (Sokal and Rohlf, 2011). Therefore, we used randomization tests based on 10,000 iterations and $\alpha = 0.05$ to estimate p values for the chi-square tests. For analyses that included course level, studies on a mixed majors courses (six studies) and studies for which course level was not reported (eight studies) were excluded. For analyses that included assessment type or inquiry type, we excluded studies that were coded as "Other" because the number of studies was small, and the assessment types or inquiry types included in the "Other" categories were quite variable. All analyses were carried out in SPSS 20.

Calculation of the effect size (i.e., the relative impact) was possible for the controlled studies. For these studies, we estimated the effect size by standardizing the difference in means between control and treatment groups by dividing the SD of the control group (see the Supplemental Materials in Ruiz-Primo *et al.* [2011] for a discussion of estimation of effect size). Effect sizes for frequency data were estimated using the effect size calculator at www.campbellcollaboration.org/resources/effect_size_input.php. Because of the limited number of studies for which we were able to calculate effect size ($n = 10$), we were unable to do any further statistical analysis on whether the effects of inquiry-based learning were greatest at particular course levels, from using particular types of inquiry, or for particular types of assessment.

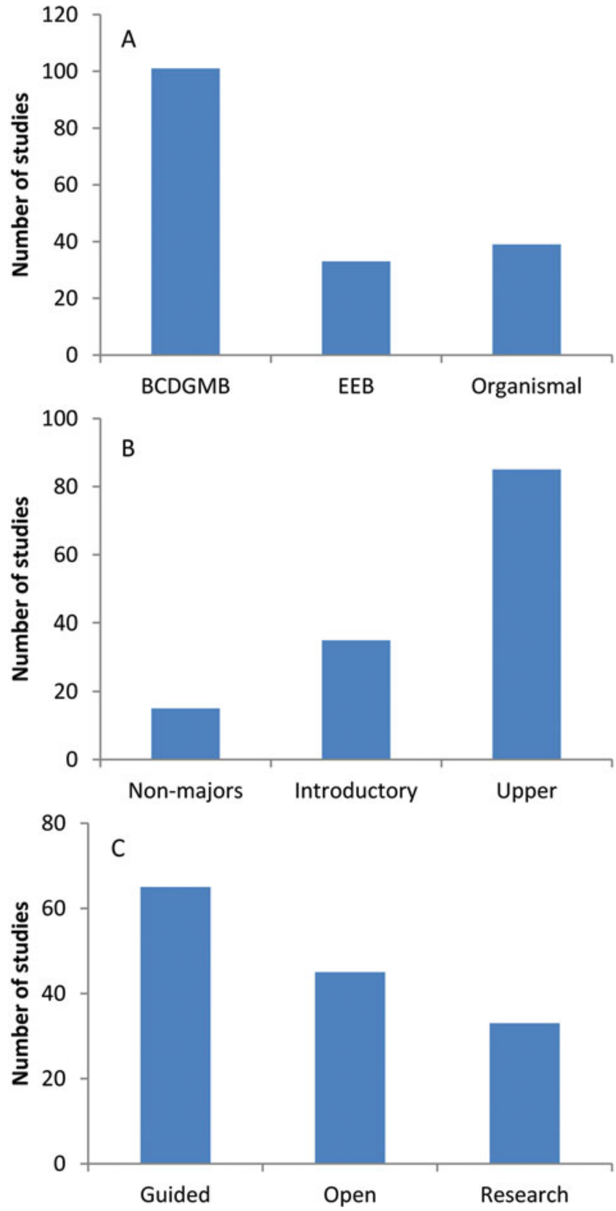

Figure 1. Number of published inquiry-based laboratory studies in biology based on (A) subdiscipline in biology: BCDGMB, biochemistry, cell biology, developmental biology, genetics, and molecular biology; EEB, ecology and evolutionary biology; Organismal, organismal biology, including plant and animal anatomy and physiology ($n = 173$); (B) course level ($n = 135$); and (C) inquiry type ($n = 143$).

RESULTS

Inquiry-Based Biology Laboratory Exercises

We found 142 papers that described inquiry-based laboratory exercises that had been implemented in undergraduate biology courses between 2005 and 2012 (see the Supplemental Material). On average, ~20 new papers were published each year. More than half of the exercises (58%) were in the areas of biochemistry, cell biology, developmental biology, genetics, and molecular biology (BCDGMB; Figure 1A), and the majority were used in courses for upper-level majors (Figure 1B). Overall, most were guided-inquiry exercises rather than open-ended exercises or research experiences imbedded in a course (Figure 1C). The approach to inquiry did not vary significantly based on course level (Table 1 and Figure 2A) or discipline (Table 1 and Figure 2B). However, BCDGMB exercises were used more often in upper-level courses and ecology and evolutionary biology (EEB) experiments were used proportionally more often in nonmajors courses (Table 1 and Figure 2C).

Assessment of Inquiry-Based Biology Laboratory Exercises

For the studies included in our sample, 73% presented assessment data (see the Supplemental Material). Of those studies that presented assessment data, almost two-thirds included

Table 1. Summary of results of cross-tabulated tests of independence[a]

Factor 1	Factor 2	p Value
Inquiry type	Course level	0.27
Inquiry type	Discipline	0.074
Discipline	Course level	<0.001*
Assessment included	Discipline	0.02*
Assessment included	Course level	0.64
Assessment included	Inquiry type	0.01*
Assessment type	Discipline	0.93
Assessment type	Course level	0.87
Assessment type	Inquiry type	0.25

[a]Rows marked with ∗ represent significant interactions between the two factors on the number of studies in our data set.

more than one type of assessment data. Qualitative assessments (such as interviews and observations) were the most frequently reported (57%), with at least one qualitative approach presented in 59 of the 104 articles that contained assessment data (see the Supplemental Material). This was

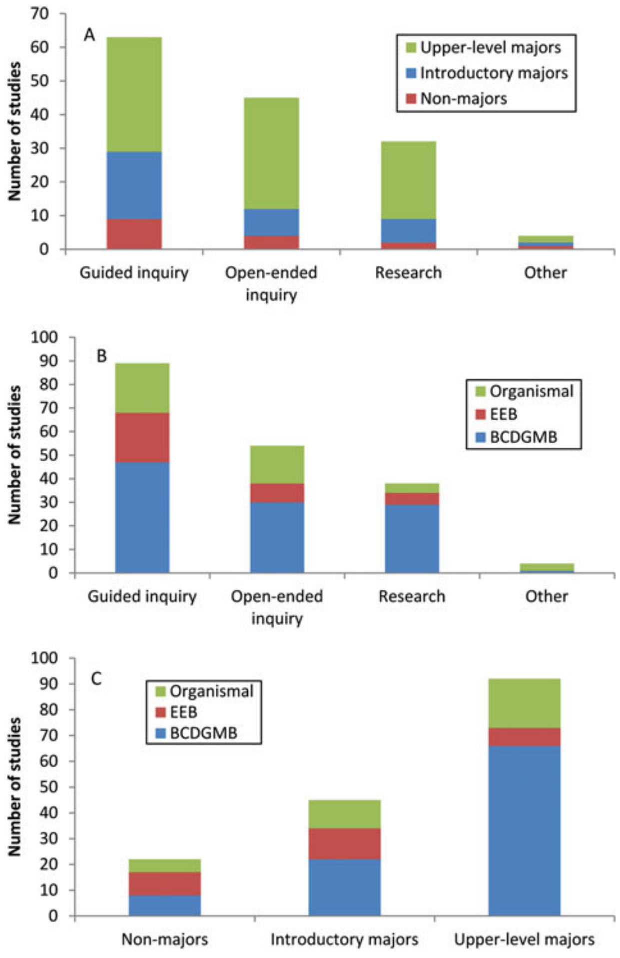

Figure 2. Number of studies based on (A) inquiry type and level at which experiment was taught ($n = 144$); (B) subdiscipline in biology and inquiry type ($n = 185$); and (C) subdiscipline in biology and level at which experiment was taught ($n = 159$).

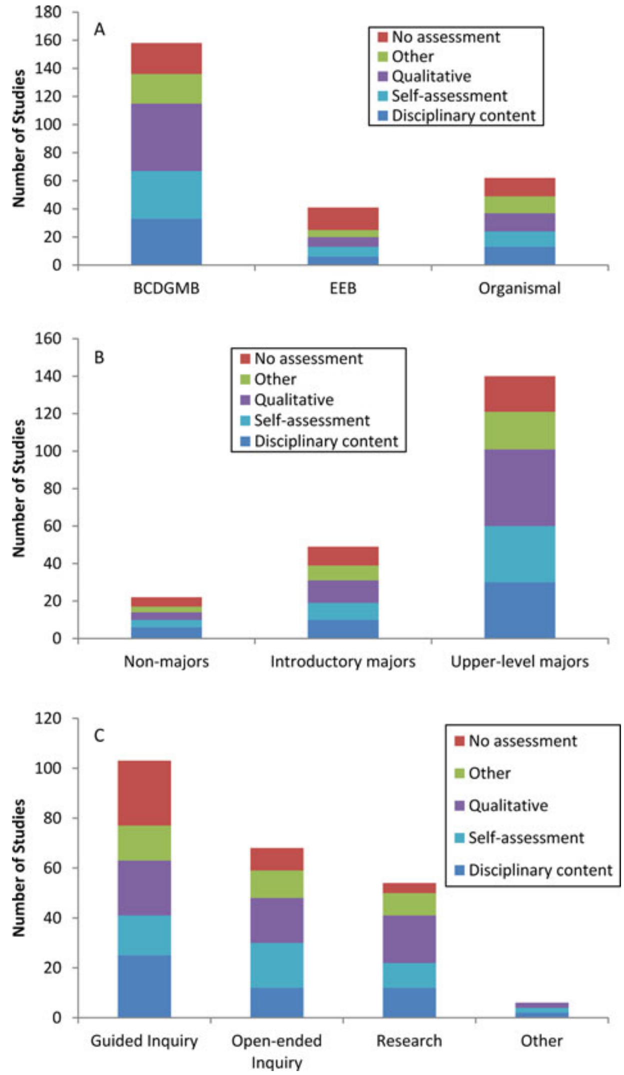

Figure 3. Number of studies based on assessment type and (A) subdiscipline in biology ($n = 261$); (B) course level ($n = 211$); and (C) inquiry-type ($n = 231$).

followed by student self-assessment and self-efficacy measures (42% of articles) and disciplinary content-type assessments like weekly quizzes or final exams (40% of the articles).

The inclusion of assessment data varied significantly based on discipline and type of inquiry (Table 1). Proportionally more BCDGMB and research exercises included assessment data, whereas proportionally fewer EEB and guided-inquiry exercises included assessment data (Figure 3, A and C). Course levels did not vary significantly in the frequency of assessment data (Table 1 and Figure 3B). For the studies with assessment data, qualitative assessment was more common than expected, assuming that the four different assessment types would occur with equal frequency ($p = 0.04$; Figure 3). Biology subdiscipline, course level, and inquiry type did not influence the type of assessment (Table 1 and Figure 3).

Although a large portion of the studies included assessment data, of those that did, only 15% were controlled

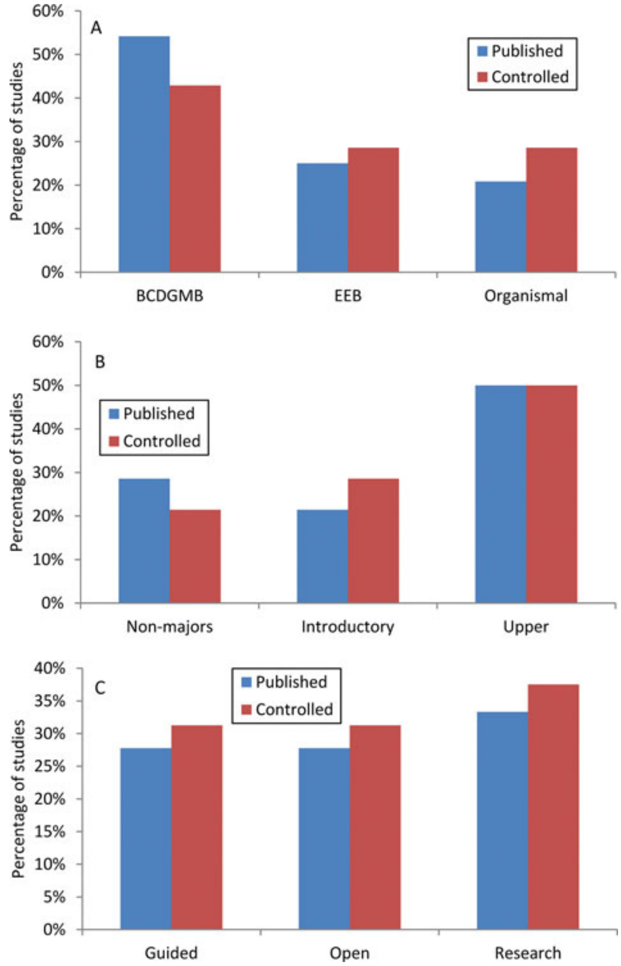

Figure 4. Percentage of studies with assessment based on published instrument (n = 17) or controlled (n = 21) divided by (A) subdiscipline in biology: BCDGMB, biochemistry, cell biology, developmental biology, genetics, and molecular biology; EEB, ecology and evolutionary biology; Organismal, organismal biology, including plant and animal anatomy and physiology; (B) course level; and (C) inquiry type.

studies and only 12% were based on published instruments. Both controlled studies and those based on published instruments were distributed based on course level and subdiscipline in biology in frequencies similar to all studies in our data set (Figures 1 and 4). Although guided-inquiry studies were more common (Figure 1C), control studies and those based on published instruments were most common in studies in which research inquiry was used (Figure 4C). Published or validated instruments that were reported include the Attitude and Intentions Survey (based on Silverstein, 1999, and AstraZeneca Science Teaching Trust, 2002), Student Assessment of Their Learning Gains (SALG; www.salgsite.org), Science Attitude Survey (Moore, 1996), and the Classroom Undergraduate Research Experience and Summer Undergraduate Research Experience surveys (www.grinnell.edu/node/25703). In other cases, studies used instruments that had been developed in the context of similar laboratory courses (Marshall, 2007; Simmons *et al.*, 2008; Dean and Wilder, 2011). For example, Dean and Wilder (2011) assessed their students who carried out a recombinant DNA exercise using a survey developed by Sleister (2007) to assess her students who isolated yeast mutants. The most commonly used published instrument, and the only instrument used in more than one study (n = 3), was the SALG. The majority of studies that used published instruments (11 out of 17 studies) were using student self-assessment. Of the remainder of the studies, one assessed disciplinary content, one used qualitative assessment, and four were classified as "other" (e.g., critical-thinking skills, scientific reasoning skills).

Most frequently, assessments covered the entire semester (61% of studies), rather than individual exercises. For papers that reported the number of students studied in their assessment, sample sizes ranged from 3 to 1018 with a median sample size of 54. Larger sample sizes were often the result of pooling data across multiple semesters. Although many studies presented assessment data for multiple semesters, relatively few presented assessment data from more than one institution (Campbell *et al.*, 2007; Belanger, 2009; Shaffer *et al.*, 2010; Baumler *et al.*, 2012). Generally, data from multiple institutions came from national projects–related biology laboratory education like the Genome Consortium for Active Teaching (Campbell *et al.*, 2007) and the Genomics Education Partnership (GEP; Shaffer *et al.*, 2010), or multi-institutional genomics projects (Baumler *et al.*, 2012). The curricula associated with these projects are predominantly used in upper-level microbiology, genetics, molecular biology, or genomics courses. Belanger (2009) used the same laboratory exercise at two different institutions where he taught.

All but one of the studies that included assessment data reported a positive effect of inquiry-based learning for at least one assessment measure, as would be expected for published experiments. In all controlled studies that measured learning gains, students in inquiry-based laboratory courses had significantly higher learning gains than students in non-inquiry courses. In addition, students in inquiry-based laboratory courses enjoyed them more and had more significant positive shifts in science attitudes as compared with students in non-inquiry laboratory courses. The one exception to the trend for a positive effect of inquiry-based learning was a study by Basey and Francis (2011). They found that students had a significantly greater preference for scripted versions of laboratory exercises as compared with inquiry-based versions of the same exercises for one of three exercises used in a first-semester introductory biology laboratory course. Students did not differ in their preference for a particular pedagogical approach for the other two exercises. Basey and Francis (2011) suggest that students might prefer scripted versions of laboratory exercises that they perceive to be more difficult.

Of the 16 controlled studies in our sample, only 10 presented sufficient data to calculate an effect size. For all assessment in those studies, the mean effect size was 0.73 (median = 0.61, SD = 0.71). When we took the average effect size for each study first, in order to avoid pseudoreplication, the mean effect size decreased to 0.62 (median = 0.40, SD = 0.76), due to several studies with multiple assessments and high effect sizes. The effect sizes seem to be greatest in introductory courses and courses that incorporate research (Table 2). In addition, gains appear to be greatest in student understanding of disciplinary content (Table 2).

Table 2. Effect size by assessment type, inquiry type, and course level[a]

	By assessment						By study					
	Sample size	Mean	Median	SD	Minimum	Maximum	Sample size	Mean	Median	SD	Minimum	Maximum
Assessment type												
Qualitative	9	0.39	0.34	0.53	−0.17	1.54	4	0.57	0.40	0.69	−0.08	1.54
Self-assessment	22	0.84	0.74	0.67	−0.15	2.79	2	0.48	0.48	0.81	−0.10	1.05
Disciplinary content	6	1.02	0.71	1.03	0.01	2.63	4	0.83	0.44	0.93	0.23	2.21
Other	3	0.33	0.27	0.34	0.02	0.70	2	0.25	0.25	0.33	0.02	0.49
Total	40	0.73	0.61	0.71	−0.17	2.79	12	0.57	0.39	0.67	−0.10	2.21
Inquiry type												
Guided inquiry	7	0.39	0.34	0.20	0.22	0.80	3	0.34	0.33	0.12	0.23	0.47
Open-ended inquiry	6	0.12	−0.05	0.49	−0.15	1.10	2	0.23	0.23	0.46	−0.10	0.56
Research	24	1.08	0.99	0.67	0.02	2.79	4	1.19	1.27	0.93	0.02	2.21
Other	3	−0.08	−0.12	0.11	−0.17	0.04	1	−0.08	−0.08	N/A	−0.08	−0.08
Total	40	0.73	0.61	0.71	−0.17	2.79	10	0.62	0.40	0.76	−0.10	2.21
Course level												
Nonmajors	7	0.36	0.34	0.24	0.02	0.80	3	0.27	0.33	0.23	0.02	0.47
Introductory majors	24	0.88	0.74	0.64	−0.17	2.79	3	0.82	0.99	0.83	−0.08	1.54
Upper-level majors	7	0.14	0.01	0.45	−0.15	1.10	3	0.23	0.23	0.33	−0.10	0.56
Not stated	2	2.21	2.21	0.59	1.80	2.63	1	2.21	2.21	N/A	2.21	2.21
Total	40	0.73	0.61	0.71	−0.17	2.79	10	0.62	0.40	0.76	−0.10	2.21

[a]Some studies included more than one assessment for which an effect size could be calculated. We averaged all assessments for each study before calculating sample statistics for the by study comparisons. Total sample size by study differs for the assessment-type analysis, because two studies included more than one assessment type for which we could calculate an effect size.

DISCUSSION

Since the publication of *BIO2010* (NRC, 2003a) and Handelsman *et al.* (2004), a large number of inquiry-based exercises for biology laboratory courses have been published in peer-reviewed, catalogued journals. Our meta-analysis revealed several interesting patterns in the development, implementation, and dissemination of inquiry-based exercises in undergraduate biology laboratory courses. First, published exercises in our data set were more likely to be in BCDGMB than EEB or organismal biology. The greater number of published exercises in BCDGMB could be due to a variety of reasons, including a greater number of laboratory courses, the ease of developing inquiry-based exercises, higher rates of curriculum development, and greater dissemination of inquiry-based labs in the peer-reviewed literature in BCDGMB. Whatever the explanation, our results suggest the need for greater emphasis on development and dissemination of inquiry-based laboratory exercises in ecology, evolutionary biology, and organismal biology.

Second, we found a greater emphasis on guided-inquiry exercises than open-ended inquiry and research-based approaches. Moving from guided-inquiry toward research-based laboratory pedagogies in undergraduate science has been shown to help retain students in science majors and to prepare them for the workplace (Weaver *et al.*, 2008). Producing students who not only have developed skills through inquiry, but who know how to carry out research, will aid in the production of more qualified postgraduate students and also provide those not going on to undertake further research with the knowledge of how science is done—this is empowering and potentially beneficial to society as a whole. Several national projects are working to advance course-based research experiences focusing on particular research projects (e.g., Howard Hughes Medical Institute [HHMI] Phage Hunters [Hanauer *et al.*, 2006]; GEP [Shaffer *et al.*, 2010]) or particular experimental approaches (e.g., Ecological Research as Education Network [Bowne *et al.*, 2011]; the Genome Consortium for Active Teaching [Campbell *et al.*, 2006, 2007]; Integrated Microbial Genomes Annotation Collaboration Toolkit [Ditty *et al.*, 2010]). In addition, the Course-Based Undergraduate Research Network (CUREnet) is facilitating collaborations among faculty members interested in incorporating research in their courses (Auchincloss *et al.*, 2014).

Finally, more exercises were aimed at upper-level courses for majors as compared with introductory courses for majors and nonmajors, with few introductory courses with research embedded included in this study. This bias toward upper-level courses could be due to the fact that more upper-level laboratory courses are taught than introductory laboratory courses or that development of inquiry-based exercises for upper-level courses might be more straightforward. However, all colleges and universities teach introductory biology laboratory courses, whereas upper-level laboratory course offerings in particular subdisciplines of biology will vary from institution to institution and may be limited. Our data suggest the need to continue the development and dissemination of inquiry-based laboratory curricula for introductory biology laboratory courses.

While we understand the constraints that faculty members face when moving toward teaching methods that require greater mentorship of students, in order to fulfill the goal of the *Vision and Change* report (AAAS, 2011) of involving *all* students in research, we will need to envision how to incorporate open-ended and research-based approaches in courses to a greater degree, especially in nonmajors and introductory majors courses. Although there are clearly more impediments to teaching in this way with large introductory or nonmajors courses, the value in terms of learning outcomes is worth the money and the effort. For example, in a large first-year biology course, Burke da Silva (2012) found that not only can first-year students undertake research projects and produce quality outcomes, but even nonmajors find value in research projects. Similarly, Harrison *et al.* (2011) showed that introductory biology students who engaged in course-based research improved their understanding of how research is done and changed their views on potential career choices.

The NRC report on assessing teaching and learning (NRC, 2003b) and Handelsman *et al.* (2004) specifically indicated the need for an increase in the assessment of teaching methods to allow more thorough determinations of their role in student learning. The vast majority of studies in our data set included some type of assessment data, and most studies with assessment data included multiple forms of assessment data. The large proportion of studies with assessment data could be due to several reasons. Biology faculty members could be commonly assessing their teaching and student learning. In contrast, this high level of assessment might be more of an artifact of the requirements for publication.

Although most studies included assessment data, few used published, validated instruments for assessing their students' learning. The SALG was the only instrument that was used in more than one study. Furthermore, the majority of studies using published instruments were using student self-assessment. However, self-reported learning gains are not necessarily correlated with actual learning gains (Falchikov and Boud, 1989). Given the availability of published, validated instruments for assessing scientific reasoning (Lawson, 1978), experimental design (Sirum and Humberg, 2011), and understanding of the nature of science (reviewed by Lederman, 2007), and the recent development of concept inventories and diagnostic question clusters in biology (D'Avanzo, 2008; Smith and Tanner, 2010), biology faculty members could be using published instruments to assess their students to a greater degree. The use of published, validated instruments will improve our ability to compare the effectiveness of inquiry-based learning in the laboratory across courses and institutions.

Most studies in our data set were limited to individual courses at a single institution. If the results of these studies can be generalized across courses and institutions remains unclear due to differences in student demographics and faculty implementation of the curricula (NRC, 2012). In the three studies that assessed students across different institutions (Campbell *et al.*, 2007; Shaffer *et al.*, 2010; Baumler *et al.*, 2012), student data were pooled across all institutions. As a result, we were unable to determine whether course or institutional contexts had an effect on student learning gains. However, a recently published study by Shaffer *et al.* (2014) suggests that differences across institutions do not appear to significantly affect learning gains.

Although many inquiry-based laboratory exercises will not be implemented in different courses at different institutions, comparisons across studies can be facilitated by the reporting of either effect sizes (e.g., Simmons *et al.*, 2008) or complete summary statistics (Ruiz-Primo *et al.*, 2011). In the controlled studies for which we were unable to estimate effect sizes, studies did not include estimates of variation (SD or SEM), did not indicate whether error bars represented SDs or SEMs, or did not provide sample sizes that are necessary to calculate SDs from SEMs. Both means and SDs are needed to estimate effect sizes, and sample sizes are also needed to estimate weighted effect sizes.

Clearly, controlled studies allow us to draw stronger inferences about the impact of inquiry-based laboratory courses than studies that do not include a control (Handelsman *et al.*, 2004; Ruiz-Primo *et al.*, 2011). In our data set, only 15% of the studies included a control group (i.e., a group of students who did not conduct the inquiry-based laboratory exercises). However, we caution that if a priori we think that a pedagogical change will improve student learning, withholding that innovation from a group of students would be unethical. In these cases, collecting data on students in semesters before a change in teaching approach and then on students in semesters after the change would balance ethical concerns with the need for experimental rigor, particularly if the assessment is done using published instruments. Also, this approach is the only tenable one for faculty members who teach at smaller institutions where only a single section of a course is taught at any particular time.

Not surprisingly, the majority of studies in our data set supports the idea that students in inquiry-based laboratory courses make significantly greater learning gains than students in non-inquiry courses. In the few studies in our data set that have sufficient information to calculate effect size, the mean effect size was 0.73. In other words, the outcome of the average student in an inquiry-based laboratory course was 0.73 SDs better than the outcome of the average student in a non-inquiry course. This effect size is greater than the effect size for a wide variety of course innovations in biology courses (Ruiz-Primo *et al.*, 2011) and is well above the threshold of 0.25 that is used as the criterion for an educational innovation to be "substantively important" (Institute of Education Sciences, 2013). However, we would caution that this large effect size also might be indicative of a file-drawer effect in which studies that show no effect or negative effects of inquiry-based learning are not submitted for publication. In addition, effect sizes were quite variable across studies (Table 2). The greatest effects seem to be found in introductory courses and courses that incorporate research (Table 2). Students also appear to show the largest gains in disciplinary content knowledge (Table 2). However, the number of studies in each category for which we were able to estimate effect sizes is very small. Many more data are needed before we can draw any conclusions as to which students are benefited most by inquiry-based learning in laboratory courses, which approaches to laboratory instruction are most beneficial, and where student learning gains are most pronounced.

Overall, our meta-analysis of the literature suggests that inquiry-based learning in biology laboratory courses can increase student learning gains. However, whether the impact of individual innovations applies across institution types and course levels is unclear. With so much evidence supporting inquiry-based laboratory teaching and its enhancement of student learning, the question remains as to why inquiry and research-based pedagogies are not found more often in university science courses. The reform is moving slowly, and this is likely due to time for faculty to develop new laboratory exercises (Spell *et al.*, 2014) or potentially due to the conflict between teaching and research in terms of professional identity for faculty (Brownell and Tanner, 2012). With the move toward teaching and education—focused faculty members who carry out more pedagogical research and who spread their knowledge to more research-focused faculty (Bush *et al.*, 2013)—perhaps the speed of implementation may soon increase.

In summary, the biology education community has made strides in developing, implementing, and disseminating new inquiry-based laboratory curricula for undergraduate courses. In addition, the community has taken beginning steps in assessing the effectiveness of these curricular changes. However, substantial gaps still exist in curriculum development and assessment that will need to be filled to meet the grade. We need to redouble our efforts in curriculum development and dissemination, especially with respect to research-based curriculum at the introductory level. Moreover, systematic assessment of current and future inquiry-based laboratory curricula across institutions using validated instruments is highly recommended.

REFERENCES

American Association for the Advancement of Science (2011). Vision and Change: A Call to Action, Washington, DC.

AstraZeneca Science Teaching Trust (2002). Attitudes to Science Teaching Questionnaire, Leicester City, UK: Leicester City Primary Science Project.

Auchincloss LC *et al.* (2014). Assessment of course-based undergraduate research experiences: a meeting report. CBE Life Sci Educ *13*, 29–40.

Basey J, Francis C (2011). Design of inquiry-oriented science labs: impacts of student attitudes. Res Sci Tech Educ *29*, 241–255.

Baumler DJ, Banta LM, Hung KF, Schwarz JA, Cabot EL, Glasner JD, Perna NT (2012). Using comparative genomics for inquiry-based learning to dissect virulence of *Escherichia coli* O157:H7 and *Yersinia pestis*. CBE Life Sci Educ *11*, 81–93.

Belanger KD (2009). Using affinity chromatography to investigate novel protein-protein interactions in an undergraduate cell and molecular biology lab course. CBE Life Sci Educ *8*, 214–225.

Bowne DR *et al.* (2011). Transforming ecological science at primarily undergraduate institutions through collaborative networks. BioScience *61*, 386–392.

Brownell SE, Tanner KD (2012). Barriers to faculty pedagogical change: lack of training, time, incentives, and tensions with professional identity. CBE Life Sci Educ *11*, 339–346.

Burke da Silva K (2012). Evolution centered teaching of biology. Ann Rev Genom Hum Genet *13*, 363–380.

Bush SD, Pelaez NJ, Rudd JA, Stevens MT, Tanner KD, Williams KS (2013). Widespread distribution and unexpected variation among science faculty with education specialties (SFES) across the United States. Proc Natl Acad Sci USA *110*, 7170–7175.

Campbell AM, Eckdahl TT, Fowlks E, Heyer LJ, Hoopes LLM, Ledbetter ML, Rosenwald AG (2006). Genome Consortium for Active Teaching (GCAT). Science *311*, 1103–1104.

Campbell AM, Ledbetter MLS, Hoopes LLM, Eckdahl TT, Heyer LJ, Rosenwald A, Fowlks E, Tonidandel S, Bucholtz B, Gottfried G (2007). Genome Consortium for Active Teaching: meeting the goals of *BIO2010*. CBE Life Sci Educ *6*, 109–118.

D'Avanzo C (1996). Three ways to teach ecology labs by inquiry: guided, open-ended, and teacher-collaborative. Bull Ecol Soc Am *77*, 92–93.

D'Avanzo C (2008). Biology Concept Inventories: overview, status, and next steps. BioScience *58*, 1079–1085.

Dean DM, Wilder JA (2011). The "Frankenplasmid" lab: an investigative exercise for teaching recombinant DNA methods. Biochem Mol Biol Educ *39*, 376–383.

Ditty JL *et al.* (2010). Incorporating genomics and bioinformatics across the life sciences curriculum. PLoS Biol *8*, e1000448.

Falchikov N, Boud D (1989). Student self-assessment in higher education: a meta-analysis. Rev Educ Res *59*, 395–430.

Hanauer DI, Jacobs-Sera D, Pedulla ML, Cresawn SG, Hendrix RW, Hatfull GF (2006). Teaching scientific inquiry. Science *314*, 1880–1881.

Handelsman J *et al.* (2004). Scientific teaching. Science *304*, 521–522.

Harrison M, Dunbar D, Ratmansky L, Boyd K, Lopatto D (2011). Classroom-based science research at the introductory level: changes in career choices and attitude. CBE Life Sci Educ *10*, 279–286.

Institute of Education Sciences (2013). What Works Clearinghouse: Procedures and Standards Handbook, version 3.0, Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Lawson AE (1978). The development and validation of a classroom test of formal reasoning. J Res Sci Teach *15*, 11–24.

Lederman NG (2007). Nature of science: past, present, and future. In: Handbook of Research on Science Education, ed. SK Abell and NG Lederman, New York: Routledge, 831–880.

Marshall PA (2007). Using *Saccharomyces cerevisiae* to test the mutagenicity of household compounds: an open ended hypothesis-driven teaching lab. CBE Life Sci Educ *6*, 307–315.

Moore RW (1996). A Scientific Attitude Inventory, Miami, OH: College of Education, Miami University.

National Research Council (NRC) (2012). Discipline-based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering, Washington, DC: National Academies Press.

NRC (2003a). BIO 2010: Transforming Undergraduate Education for Future Research Biologists, Washington, DC: National Academies Press.

NRC (2003b). Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics: Report of a Workshop, Washington, DC: National Academies Press.

Ruiz-Primo MA, Briggs D, Iverson H, Talbot R, Shepard LA (2011). Impact of undergraduate science course innovations on learning. Science *331*, 1269–1270.

Shaffer CD *et al.* (2010). The Genomics Education Partnership: successful integration of research into laboratory classes at a diverse group of undergraduate institutions. CBE Life Sci Educ *9*, 55–69.

Shaffer CD *et al.* (2014). A course-based research experience: how benefits change with increased investment in instructional time. CBE Life Sci Educ *13*, 111–130.

Silverstein S (1999). Preprogram Survey of the Scientific Work Experience Program for Teachers (SWEPT), New York: Columbia University.

Simmons ME, Wu XB, Knight SL, Lopez RR (2008). Assessing the influence of field- and GIS-based inquiry on student attitude and conceptual knowledge in an undergraduate ecology lab. CBE Life Sci Educ *7*, 338–345.

Sirum K, Humburg J (2011). The Experimental Design Ability Test (EDAT). Bioscene *37*, 8–16.

Sleister HM (2007). Isolation and characterization of *Saccharomyces cerevisiae* mutants defective in chromosome transmission in an undergraduate genetics research course. Genetics *177*, 677–688.

Smith JI, Tanner K (2010). The problem of revealing how students think: concept inventories and beyond. CBE Life Sci Educ *9*, 1–5.

Sokal R, Rohlf F (2011). Biometry, 4th ed., New York: Freeman.

Spell RM, Guinan JA, Miller KR, Beck CW (2014). Redefining authentic research experiences in introductory biology laboratories and barriers to their implementation. CBE Life Sci Educ *13*, 102–110.

Sundberg MD, Armstrong JE (1993). The status of laboratory instruction for introductory biology in U.S. universities. Am Biol Teach *55*, 144–146.

Sundberg MD, Armstrong JE, Wischusen EW (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges & universities. Am Biol Teach *67*, 525–529.

Tuss P, Frechtling J, Ewing T, Westat I, Sladek M, McBride D (1998). A Report on an Evaluation of the National Science Foundation's Instrumentation and Laboratory Improvement Program, Washington, DC: National Science Foundation.

Weaver GC, Russell CB, Wink DJ (2008). Inquiry-based and research-based laboratory pedagogies in undergraduate science. Nat Chem Biol *4*, 577–580.