

The Effect of a Paired Lab on Course Completion and Grades in Nonmajors Introductory Biology

Dayna Jean DeFeo,^{1*} Andrew Bibler,² and Sarah Gerken³

¹Center for Alaska Education Policy Research, ²Institute of Social and Economic Research, and

³Biological Sciences Department, University of Alaska Anchorage, Anchorage, AK 99508

ABSTRACT

This paper explores the effect of a paired lab course on students' course outcomes in nonmajors introductory biology at the University of Alaska Anchorage. We compare course completion and final grades for 10,793 students (3736 who simultaneously enrolled in the lab and 7057 who did not). Unconditionally, students who self-select into the lab are more likely to complete the course and to earn a higher grade than students who do not. However, when we condition on observable course, academic, and demographic characteristics, we find much of this difference in student performance outcomes is attributable to selection bias, rather than an effect of the lab itself. The data and discussion challenge the misconception that labs serve as recitations for lecture content, noting that the learning objectives of science labs should be more clearly articulated and assessed independent of lecture course outcomes.

This paper explores the effect of a paired lab course on students' course outcomes in introductory biology for nonmajors at the University of Alaska Anchorage (UAA), a large, open-enrollment, 4-year university. We compare outcomes for 10,793 students, 3736 who simultaneously enrolled in the lab and 7057 who did not, and analyze the degree to which they select into the lab on observable characteristics to explore the following research questions:

1. Are students who take a paired lab more likely to complete the lecture component (i.e., receive a final grade as opposed to withdrawing or receiving an Incomplete)?
2. Are students who take a paired lab more likely to receive a higher grade in the lecture component?
3. Does the laboratory experience differently affect course outcomes for students in specific demographic subgroups (e.g., gender, race, high school urbanicity, age, prior academic performance, and socioeconomic status)?

HISTORY OF SCIENCE LABS

Secondary and postsecondary science classes traditionally have two components: the lecture and the laboratory. A learning experience fundamentally different from the classroom, the lab offers the hands-on opportunity for students to “interact with materials and/or with models” (Hofstein and Lunetta, 2004, p. 31). Though labs are germane to science education, their history in the curriculum reveals that, although their purposes and learning objectives have changed fundamentally since their inception, the structure of separate labs and lectures has held relatively steady.

The first science labs in higher education were offered in England in the discipline of chemistry in 1807 (Reid and Shah, 2007), and the United States quickly followed suit (Blosser, 1983); the objective of these early lab experiences was to train students

Elisabeth Schussler, *Monitoring Editor*

Submitted Mar 9, 2020; Revised May 8, 2020;

Accepted Jun 16, 2020

CBE Life Sci Educ September 1, 2020 19:ar36

DOI:10.1187/cbe.20-03-0041

*Address correspondence to: Dayna Jean DeFeo (ddefeo@alaska.edu).

© 2020 D. J. DeFeo et al. CBE—Life Sciences Education © 2020 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–Noncommercial–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.

for careers as technicians or bench scientists. In 1878, the U.S. Commissioner of Education advocated for lab work in all physical science classes, and Harvard University's decision to require high school physics labs as an admission requirement in 1886 heavily influenced how secondary schools delivered science education (Blosser, 1983). In 1893, the National Education Association Committee of Ten called for science education—including labs—for all secondary students, therein affirming the place of science in the general education curriculum, which had previously emphasized the humanities (Tamir, 1976; DeBoer, 2000), and secondary and postsecondary science incorporated labs even at the introductory levels.

At the turn of the century, Armstrong's work in the United Kingdom (see Klainin, 1988) and Dewey's work in the United States spurred the progressive education movement, and constructivist approaches reframed labs (Hofstein and Lunetta, 1982) as opportunities for students to connect learning and schema. As Dewey emphasized education's role of preparing students to *participate* in a democratic society (as opposed to serving its *economic* needs), a perennial question in our field was first posed in the 1920s (DeBoer, 2000): Was the objective of science education to support future citizens to apply the general principles of science, or was it to support the development of future scientists (Bybee, 2015)?

Through World War II and the subsequent launch of Sputnik, the national agenda for science education remained utilitarian: to serve national security and infrastructure and economic growth objectives (Rudolph, 2014), but concomitantly, the Cold War brought newfound attention to the social responsibility that accompanies science and technological innovations (DeBoer, 2000). The concept of science literacy was introduced in the 1950s (Holbrook and Rannikmae, 2009; Bybee, 2015). Rather than discrete content knowledge, science literacy emphasized applying scientific methods and critical-thinking skills. This was a fundamental shift in learning objectives for science education and the pedagogical approaches applied in labs—which up to that point had largely been demonstration based or confirmatory (“cookbook” style) in nature (Hofstein and Lunetta, 1982; Klainin, 1988).

Influenced by Bruner and Piaget (Hofstein and Lunetta, 2004), science education underwent major curricular reforms in the 1960s (Klainin, 1988; Hofstein, 2004), and the lab was positioned as the center of science education (Hofstein, 2004). The focus shifted from confirmatory activities to developing critical thinking through interaction with materials, with the lab being a place where students could understand the nature of science (Hodson, 1993). In the 1980s, Vygotsky's constructivist theory gained traction, and science education in the United States emphasized inquiry-based approaches as more authentic models of the scientific process (Hodson, 1988; Hodson and Hodson, 1998; Hofstein and Lunetta, 2004). In 1996, the National Science Education Standards from the National Research Council (NRC) declared that inquiry should be a part of science education at all levels, and this position has consistently been reaffirmed (NRC, 2000; American Association for the Advancement of Science, 2011). Curricular revisions to shift secondary and postsecondary science labs to inquiry-based approaches were heavily supported in the early 2000s through federal and private funding (see Gess-Newsome et al., 2003). However, science departments must consider not only how labs

are delivered, but the effect of these activities on student learning (Beck et al., 2014).

LEARNING OUTCOMES FOR THE SCIENCE LAB

From this history emerges a contemporary challenge to studying the effectiveness of science labs: their “plurality of purposes” (Reid and Shah, 2007, p. 177). Though the discourse setters in the field speak in broad terms about science literacy as an overarching goal, discrete student learning outcomes for general science labs are much more variable. The task of distilling learning outcomes for science labs has been undertaken several times (Tamir, 1976; Kirschner and Meester, 1988; Reid and Shah, 2007; see also Garnett et al., 1995, for a listing of previous investigations through the early 1990s). Our review of literature published since the 1990s identified five categories of learning objectives that, incidentally, are generally aligned with previous analyses that have considered these outcomes across science disciplines: 1) *promote deeper critical thinking or problem-solving skills*, including collaboration and communication (NRC, 2000; Hofstein and Mamlok-Naaman, 2007; Momsen et al., 2010); 2) *develop content knowledge*, including the joining of theory with application (Bybee, 2000; Carnduff and Reid, 2003; Reid and Shah, 2007); 3) *apply scientific methods and understand the nature of science* (Adams, 1998; Krajcik et al., 2001; Hofstein and Mamlok-Naaman, 2007; Reid and Shah, 2007; Beck and Blumer, 2012); 4) *develop practical and technical skills* (Bybee, 2000; Howard and Miskowski, 2005; Reid and Shah, 2007; see also Carnduff and Reid, 2003, who emphasize safety as a component of this objective); and 5) *develop positive attitudes*, including confidence, interest, and engagement in science (Carnduff and Reid, 2003; Hofstein and Lunetta, 2004; Reid and Shah, 2007).

Our analysis focuses on content knowledge outcomes. Though labs have been a cornerstone of the science curriculum for more than a century, there is relatively little research on their efficacy in this area (see also Hofstein and Lunetta, 1982, 2004, and Tobin, 1990, who lament a dearth of empirical studies on labs as a perennial challenge in the science education discipline). Though the recent empirical literature on science labs has focused more on affective outcomes, content knowledge is the most commonly reported student learning outcome in the practitioner literature (Puttick et al., 2015) and the foremost goal of the science curriculum during the accountability movement. Recent studies have documented that active-learning strategies used in the lab lead to better test performance (Thompson and Soyibo, 2002; Howard and Miskowski, 2005) and student perceptions of their own knowledge gains (Browne and Blackburn, 1999; Adami, 2006; Howard and Miskowski's, 2005; Kloser et al., 2013). Hart (2018) noted that *faculty presence* in the lab (as opposed to teaching assistants) and *better-structured assignments* promoted content knowledge gains, as indicated by final exam scores and final grades. Recently, research on labs' support of content knowledge outcomes has also considered the efficacy of virtual or online labs and measured these outcomes through final grades or content exams (Son et al., 2016; McQueen, 2017).

Though studies generally detect positive learning outcomes, Puttick et al. (2015) lament that this scholarship suffers from “innovation bias,” that is, “researchers focusing on novel interventions instead of common or standard curricula” (p. 13). We

also note that many studies are limited by small sample sizes and a lack of randomization that limits both the complexity of their analyses and the strength of their conclusions (see also Hofstein and Lunetta, 2004; Brownell *et al.*, 2013; Beck *et al.*, 2014). We empathize with the researchers: It is not easy to do randomized control trials with students (Grossman and Mackenzie, 2005; Sullivan, 2011) or to ensure implementation fidelity around new teaching methods (Meyers and Brandt, 2014), and the majority of studies are conducted by faculty members who are themselves teaching courses and are unable to control for many of the variables.

However, our interest is not only study quality, but the focus of these investigations. Recent studies explore not the lab experience in general, but a specific pedagogy applied within the lab environment. When comparisons are used, they are generally between “good teaching” and “bad teaching,” but both groups receive the treatment of the lab experience. Thus, our concern with previous investigations is around the obvious confounding variable validity threat: The content knowledge goals for the lab are similar to the goals for the accompanying science course (Hofstein and Lunetta, 1982), and as labs are almost always paired with lecture, student learning gains could be the result of either treatment.

Matz and colleagues’ (2012) study took advantage of non-concurrent chemistry lab and lecture enrollment on their campus to identify a correlation between lab enrollment and lecture performance. They identified a correlation outright and, after controlling for high school grade point average (GPA), Scholastic Aptitude Test scores, and science placement exam scores, the strength of the correlation decreased, but still existed, and was strongest for students in the lowest-performing academic quartile. However, their interpretation that the residual effect is causal is problematic, as the analysis omits other variables likely to affect performance, such as instructor quality, nonacademic demographic student characteristics, and student motivation. We apply a similar approach but include more variables in the regression to see how the strength of the correlation holds when we more precisely condition the analysis.

SETTING AND DATA

At UAA, undergraduate students who are not science majors take two introductory-level science classes as part of their general education requirements, but they are required to take only one lab. Thus, the focus of our analysis, Biology 102: Fundamentals of Biology (BIOL 102), a course for nonmajors, can be taken alone as a lecture class or paired with a 1-credit lab course, Biology 103: Fundamentals of Biology Laboratory (BIOL 103). In each lecture section, some students are concurrently enrolled in the paired lab, while others are not. Though some institutions are shifting away from concurrent enrollment in lecture and lab (see Matz *et al.*, 2012), lab enrollment is typically mandatory; thus, the unusual opt-in situation at UAA provides an opportunity to study the extent to which selection on observable characteristics into the lab drives the correlation between lab enrollment on course outcomes. Although the proposed analysis falls short of comparisons between randomly assigned treatment and control groups, the large and diverse data set allows us to condition on a variety of student and course variables to learn about self-selection into the lab.

Because of the documented extreme variability in the way science labs are administered, Hofstein and Mamlok-Naaman (2007) recommend that research about labs be explicit in describing the experience there (see also Lunetta *et al.*, 2007). Thus, to provide necessary background and context for the analysis, we describe the institution and student body; the BIOL 102 lecture course structure, content, and grading system; and the BIOL 103 lab course content and structure.

Institution and Student Body

UAA is a midsized (14,955 undergraduate students in 2018) open-enrollment institution offering graduate and undergraduate programs. The city of Anchorage is rich in ethnoracial diversity (Farrell, 2016), which is reflected in enrollments; 41% of UAA students identify as non-White. Attendance patterns and academic preparation at UAA are typical to open-enrollment institutions: 54% of students attend part-time and 35% receive Pell Grants (UAA Office of Institutional Research, 2015); 63% require developmental education in at least one area (University of Alaska Statewide Office of Institutional Research, 2016).

Dependent Variable: Lecture Course Outcomes

BIOL 102 at UAA is a fairly typical biology survey course; the lecture is always taught by a tenured, tenure-track, or term faculty member with a PhD in biology or a related field. The course is also managed by a faculty coordinator, who ensures alignment and consistency of content and assignments across sections. The learning objectives are typical for introductory non-majors biology and include applying scientific methods, concept knowledge, history and major advances in the discipline, and applying biological concepts to understand natural processes. At UAA, biology and astronomy are the only two general education sciences without a math prerequisite, and thus, relative to other introductory science courses, BIOL 102 disproportionately enrolls first-year students and students who are concurrently taking developmental math.

We conducted a syllabus review of all sections of BIOL 102 included in our analysis ($n = 291$) to ascertain course structure, content, and grading criteria. During the period of analysis, the class was largely organized around an introductory biology textbook¹ and associated publisher-developed homework assignments. Face-to-face classes met for 75-minute class sessions twice per week, and online offerings used commensurate asynchronous recorded lectures. Student time in class was spent largely in standard lecture with occasional clicker questions, with few (if any) group- or team-based activities; online lectures included required student responses (via clicker-style questions) and in some instances required participation on discussion boards via the course management system.

BIOL 102 final course grades are the dependent variable for our investigation, and comprised primarily multiple-choice exams (three per semester), automatically graded publisher-developed online homework assignments, and weekly quizzes. Though weekly quizzes reflect both attendance and content knowledge (as points could not be made up for students who did not attend class), the other assessments (and the final grade

¹Textbooks for the class varied during the period of investigation, but most courses used Hoefnagels (2015a,b) *Biology: The Essentials* or Mader and Windelspecht (2014, 2017a,b), *Essentials of Biology*.

derived from them) largely measured discrete content knowledge. There was some variation of point distribution between years, but the majority of points were earned on the three content knowledge exams. The focus on discrete biological concepts in spite of stated course goals of science literacy is fairly typical; Bybee (2015) notes that, in spite of curricular advances, most schools still emphasize and assess content knowledge.

Independent Variable: Lab Enrollment

During the period of analysis, BIOL 103 met for 2 hours and 45 minutes once per week, and students completed a prelab reading and activity before the lab as a preparatory exercise. These activities were guided by an institution-developed lab manual that largely consisted of worksheets for each lab activity. The lab was taught by term or adjunct faculty (with MS or PhD degrees) and graduate student teaching assistants, supervised by the course coordinator. Class typically started with a short lecture (20–30 minutes) with background information and instruction for the lab, followed by time for students to work on group or paired projects in which they defined their own questions, proposed hypotheses, created methods, tested them, and delivered written papers and presentations (either as talks or a posters). Though not explicitly stated, there was (and remains) a general expectation within the department for the lab to support the lecture content. Thus, although we acknowledge that lab outcomes are discrete from lecture course performance, we expected to detect a correlation.

Data Source: Institutional Records

Our analysis uses UAA institutional data on course enrollments and outcomes and student-level demographics. We restricted the sample to students enrolled in their first attempt in an introduction to biology course (BIOL 102) between the Fall of 2008 and Summer of 2018 and completed the course with a valid letter grade or received an Incomplete or Withdrawal from the course. We do, however, discard observations of students who audited, received a No Basis grade,² or who were repeating the course from a previous attempt. Our two main outcomes were whether the student completed the course and the final grade the student received, and our analysis compares those outcomes between students who did and did not enroll in the lab.

Figure 1 shows the distribution of students based on whether or not they enrolled in the lab, that is, whether they are considered *treated*, and demonstrates the general comparison between the treatment (Lab) and control (No Lab) groups. Ultimately, the analysis included 3736 students in the treatment group, and 7057 in the control. The first outcome, “DNF,” refers to the proportion of students who did not finish the course, meaning they received an Incomplete or withdrew. Figure 1 illustrates that the proportion of Lab students who did not finish the course is about 8.5%, which is 4.5 percentage points lower than the proportion of No Lab students, 13.0%.

In addition to the difference in course completion rates, Figure 1 displays the proportion of Lab and No Lab students receiving each final letter grade (“A”–“F”). About 32% of Lab students received an “A” in the course, compared with 24% of the No Lab students. Similarly, Lab students were slightly more

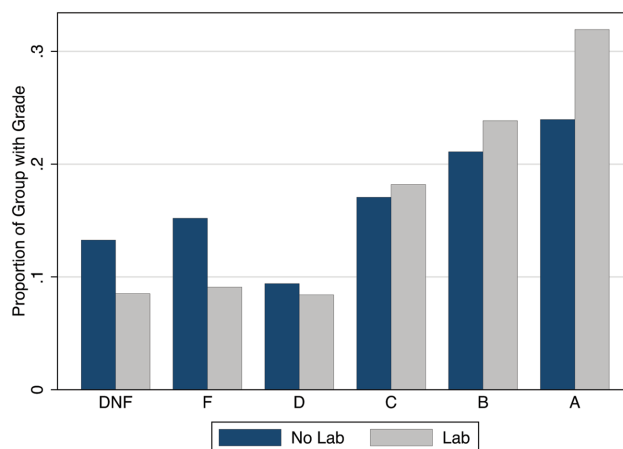


FIGURE 1. Grade distribution by lab enrollment. Unconditionally, Lab students are about 5% more likely to complete the lecture course, and 5% less likely to receive an “F”.

likely to receive a “B”, and about 5% less likely to receive an “F”. While the unconditional differences displayed in Figure 1 represent a good starting point, the crux of our analysis considers the extent to which these differences may be attributable to selection bias as we condition on student and course characteristics.

Table 1 displays the sample averages for selected outcomes and covariates, including the main outcomes of our analysis. Each column displays sample averages with standard deviations in parentheses for Lab and No Lab students. Row 2 (Grade) compares BIOL 102 letter grades between Lab and No Lab students (0–4 scale), and notes that the Lab students scored a 2.7 on average, compared with an average of 2.3 among No Lab students. While these unconditional correlations tell us that Lab students are more likely to complete the course and score higher grades, there could be many reasons for this. For example, Table 1 notes that Lab students have a higher average cumulative GPA than No Lab students, and these higher-performing students may be more likely to complete the course and score a higher grade, regardless of the lab experience. This sort of confounding factor would suggest that the unconditional correlation shown in Figure 1 and Table 1 tends to overestimate the effect of lab enrollment on course completion and grades.

Our strategy to account for some of these differences and move toward a causal interpretation is to track the change in the correlation between taking the lab and BIOL 102 course outcomes as we condition on other observable characteristics. In addition to outcome data, we grouped the students by lecture section. This is important for methodological purposes, as class-specific influences, such as the time of day that the lecture meets and the instructor’s quality, enthusiasm, or preparation, are potential confounding factors that could influence students’ performance in the course. Classroom-level information is displayed in the second section of Table 1. No Lab students had an average lecture class size of 68.0, similar to 66.3 for Lab students. However, the proportion of a lecture enrolled in a lab is different between the two samples: 0.43 for Lab students, relative to 0.3 for No Lab students, and the likelihood of taking a lab is also related to whether the student took the BIOL 102 course in a face-to-face or online section (38% of face-to-face students took the lab, compared with 18% of online students).

²A nonacademic grade reflecting a lack of attendance or work, such that there is nothing with which to evaluate the student.

TABLE 1. Summary statistics^a

	1 Lab	2 No Lab
Outcomes		
Completed course	0.915 (0.279)	0.867 (0.339)
Grade	2.668 (1.247)	2.344 (1.332)
Grade = "A"	0.319 (0.466)	0.240 (0.427)
Grade = "B"	0.238 (0.426)	0.211 (0.408)
Classroom characteristics		
Observations in lecture	66.29 (44.93)	67.95 (42.36)
Proportion of lecture in lab	0.431 (0.177)	0.299 (0.152)
Face-to-face	0.897 (0.304)	0.755 (0.430)
Demographics		
Age	22.71 (7.024)	23.24 (7.150)
Female	0.612 (0.487)	0.582 (0.493)
White	0.594 (0.491)	0.541 (0.498)
Black	0.0375 (0.190)	0.0558 (0.230)
AK Native	0.0931 (0.291)	0.110 (0.313)
Hispanic	0.0766 (0.266)	0.0840 (0.277)
Pell eligible	0.251 (0.434)	0.276 (0.447)
Academic characteristics		
Cumulative GPA	2.800 (0.932)	2.606 (1.021)
In-term credits attempted	11.90 (3.524)	11.04 (3.675)
In-term credits earned	9.627 (4.839)	8.196 (4.976)
Cumulative credits attempted	40.62 (31.08)	40.51 (34.08)
Cumulative credits earned	33.09 (26.31)	31.43 (28.42)
Accuplacer Reading score	86.40 (13.63)	85.70 (13.90)
Accuplacer Writing score	91.25 (12.99)	90.66 (13.42)
Accuplacer Arithmetic score	66.08 (8.850)	66.11 (8.769)
Accuplacer Algebra score	66.64 (20.29)	65.75 (20.72)
Accuplacer college algebra score	38.68 (8.619)	38.53 (8.150)
ACT score	21.20 (1.770)	21.16 (1.670)
HS GPA (nonmissing)	3.017 (0.514)	2.945 (0.519)
Observations	3736	7057

Our data also contain baseline student-level demographic information, including age, race, and Pell eligibility, and academic information such as high school GPA, entrance exam scores, and college GPA. Selected variables are displayed in the third and fourth sections of Table 1. Of note, UAA accepts a variety of different tests for student placement, and not all students are required to take all the placement tests in order to enroll; this resulted in a nontrivial proportion of observations with missing test data. On the two most commonly used tests, Accuplacer Reading and Accuplacer Algebra, 54% and 55% of Lab students have scores compared with 56% and 54% of No Lab students, respectively. Lab students who took the tests scored about 1 point higher in reading and algebra than No Lab students; considering the scale of the placement tests (20–120), the 1-point difference is negligible. ACT and high school GPA were also similar between the two groups; however, college GPA and credits attempted and earned in the semester before the biology course enrollment are slightly higher for Lab students and may be indicative of selection bias. Thus, our analysis accounts for these additional classroom, demographic, and academic variables, allowing us to refine our estimate of the relationship between taking the lab and course outcomes by conditioning on observable characteristics.

METHODS

The objective of our analysis is to test the relationship between lab enrollment and lecture course outcomes in introductory biology. If lab enrollment were randomly assigned, then we could simply compare average outcomes for Lab and No Lab students, as shown in equation 1:

$$Y_{ict} = \alpha + \beta \cdot Lab_{ict} + \varepsilon_{ict} \quad (1)$$

where Y_{ict} and Lab_{ict} represent the outcome of interest, course completion or grade, and a dummy variable for enrolling in the lab, respectively, for student i enrolled in lecture c at time t . Under random assignment, our estimate for β based on equation 1 would be a valid estimate of the causal effect of taking the lab on lecture outcomes. We start by comparing course completion and lecture grades of Lab and No Lab students as in equation 1, and Figure 1 and Table 1 show there is a positive correlation between lab enrollment and course completion and final grade.

While we use these estimates, which are based on equation 1, as a starting point for our analysis, the main concern with interpreting the unconditional correlation as a causal relationship is that students self-select into the lab. Students who decide to enroll in the lab are likely different in observable and unobservable ways that are related to lab enrollment and course outcomes. In other words, ε_{ict} from equation 1 is correlated with both the decision to enroll in the lab and Y_{ict} , and the resulting estimates thus suffer from omitted variable bias, which precludes a causal interpretation. Establishing a causal relationship, or lack of, requires a more thorough analysis.

^aBecause students self-select into the lab, our analysis considered observable classroom, academic, and demographic characteristics. Additional observable student characteristics not displayed in this table, but included in our analysis are high school urbanicity (rural/urban), degree type (e.g., AA, AAS, BA, BS), and academic major.

With this empirical issue in mind, we further analyze the relationship between lab enrollment and course outcomes by comparing the unconditional correlations with conditional correlations. We apply the concept behind Altonji and colleagues' (2005) bounding strategy to analyze the stability of the coefficient of interest while adding more controls by making use of additional course, demographic, and academic information. This and similar methods have been used to assess selection bias in a number of settings, including the impact of Catholic schooling on high school graduation and college attendance (Altonji *et al.*, 2005), maternal behavior on child IQ and health (Oster, 2019), and the impact of postsecondary scholarships on college attendance and time use (Barrow and Rouse, 2018). As it applies in our case of lab enrollment and lecture performance, if controlling for observable characteristics in increasingly restrictive regressions dramatically reduces the strength of the correlation, then we can infer that selection bias is inflating the correlation. Consider equation 2, in which we separate ϵ_{ict} into several components:

$$\epsilon_{ict} = C_c + \theta^o \cdot X_{ict} + \theta^u \cdot Z_{ict} + u_{ict} \quad (2)$$

where C_c represents lecture-specific fixed effects, X_{ict} and Z_{ict} are vectors of observable and unobservable student level characteristics, respectively, and u_{ict} is an idiosyncratic error term. We compare the adjusted correlation with the initial correlation estimated in the first step using multiple regression to internally adjust the estimated correlation between the lab treatment and the lecture grade. When these characteristics are strong predictors of lab enrollment and have direct effects on course outcomes, including them in the regression will impact the estimated effect of the treatment on the outcomes, which casts doubt on a causal interpretation. We do this by comparing estimates based on equation 1 with estimates from the regression that includes more controls as based on equation 3:

$$Y_{ict} = \alpha + \beta \cdot \text{Lab}_{ict} + C_c + \theta^o \cdot X_{ict} + \theta^u \cdot Z_{ict} + u_{ict} \quad (3)$$

Using multiple regression to condition on observable characteristics develops our analysis by considering how conditioning impacts the estimated correlation, which indicates the direction of the selection bias. For example, if higher-achieving students decide to enroll in the lab simultaneously with the lecture, we would expect the correlation to shrink after conditioning on student achievement, reflecting positive selection between taking the lab and student performance. In other words, Lab students would have performed better in the lecture class even without the lab experience. If conditioning on characteristics reduces the strength of the correlation between the lab treatment and lecture grade, we should be skeptical that any remaining correlation is due to a causal effect, as it may be due to bias from other omitted variables (Z_{ict} in equation 3).

We include two refinements of our estimates based on conditioning on observable characteristics. First, we estimate the correlation while including lecture-specific fixed effects in our regression (meaning they attended a class with the same schedule and the same professor in the same semester and therefore received the same lecture treatment and course materials, had the same set of peers in the class, etc., all of which could influence the student's outcome). Limiting our comparisons in this

way avoids attributing some of the effect of lecture-specific characteristics to the lab enrollment. For example, if lenient professors also encourage their students to enroll in the lab, we might estimate a positive correlation between grades and taking the lab, but it would not actually be informative about the lab's effect on the outcome. Restricting our comparison to students in the same lecture section addresses this threat. Additionally, though our syllabus review noted that lecture classes largely covered the same content and used similar assignments to assign grades, restricting the comparison to students in the same lecture section also controls for differences in how individual instructors scored assignments (e.g., offering extra or partial credit) and ultimately assigned final grades. Thus the remaining difference between this estimate and our first one must be attributable to something other than lecture section characteristics.

Second, we include additional student-specific demographic and academic control variables (adding X_{ict}) to our regression. By restricting our comparisons even further, we refine our estimates and reduce the number of unobserved characteristics that could lead to bias. As we refine our comparisons of Lab and No Lab students, we can assess how much of the correlation between lab enrollment and lecture grade might be due to a causal relationship and how much might be due to selection bias in the nonrandom assignment. Essentially, as we move from one type of estimation (equation 1, the simple correlation), to the most-restrictive estimation method (equation 3, conditional on lecture fixed effects and all student-level covariates), if the correlation is unstable or disappears, this casts doubt that there is a causal relationship between the lab treatment and lecture performance. This possibility is shown explicitly in equations 2 and 3 by including a term to capture unobservable characteristics, Z_{ict} . The crux of this analysis is that we can assess the impact that adding C_c and X_{ict} has on our estimates. If we expect our estimates to change even further in the event that we could add Z_{ict} as well (e.g., by including data around students' motivation or attitudes about science), then we can extrapolate that an underlying true causal effect is unlikely or at least limited. On the other hand, if the estimated effect is stable across these different approaches, we can argue that at least some portion of the correlation is due to a causal effect. That is, if adding C_c and X_{ict} has no impact on our estimates, then we would not necessarily expect adding Z_{ict} to change the estimates either.

RESULTS

In Table 2 we report results for four different outcomes. The first column shows results for an outcome that is a dummy variable, indicating that the student finished the course (meaning he/she did not withdraw or have an Incomplete). The second column is the numeric version of the student's grade on a four-point scale. Columns three and four show results for two indicators that are equal to 1 if the student finished with an "A" in the course or with an "A" or "B" in the course, respectively. All coefficients in the table are from a regression of one of these outcomes on a dummy variable that indicates that the student took the lab. The corresponding standard error and the R^2 is below each coefficient estimate.

In the first row of results, "No Controls," the only regressor included was the lab dummy variable; these estimates represent differences in the unconditional averages of the outcome variable between Lab and No Lab students. The first column of

TABLE 2. Effect of taking the lab on student outcomes

	Completed course 1	Grade 2	Pr(A) 3	Pr(A or B) 4
No controls (equation 1)	0.0475*** (0.00658) 0.005	0.331*** (0.0351) 0.013	0.0797*** (0.0111) 0.007	0.107*** (0.0116) 0.010
Classroom fixed effects (equation 2)	0.0449*** (0.00634) 0.004	0.320*** (0.0308) 0.012	0.0803*** (0.00967) 0.008	0.0953*** (0.0106) 0.008
Full controls (equation 3)	0.0254*** (0.00585) 0.143	0.117*** (0.0190) 0.587	0.0355*** (0.00758) 0.279	0.0306*** (0.00800) 0.382
Number of observations	10,793	9539	10,793	10,793
Number of classrooms	291	291	291	291

*As we conditioned our analysis on more control variables, the correlation between the lab experience and lecture course outcomes diminished, which suggests that the observed differences in student performance are attributable to selection bias, and the lab itself has no causal effect on the probability of completing the course or earning a passing grade.

Pr = Probability.

the first row suggests that Lab students were 4.75 percentage points more likely to finish the course; this difference is equivalent to the one shown in Figure 1 in the “DNF” column. In the second row, the estimated coefficient is conditional on classroom fixed effects. The comparison is still between Lab and No Lab students, but is restricted to comparing students in the same lecture section. After conditioning on the lecture section, the estimated difference is almost identical to the unconditional difference, 4.5 percentage points.

Finally, in the last row of results, “Full Controls,” we add all additional demographic- and performance-related controls to the regression, so that now the estimated difference between Lab and No Lab students is conditional on the specific lecture section as well as all the available student-level control variables. Comparing this with the unconditional correlation provides some indication of how much of the unconditional gap is due to selection on observable characteristics. For the Completed Course outcome, after conditioning on all of the observable characteristics, the estimated difference is 2.5 percentage points. This is about half the size of the unconditional correlation, and suggests that the observable characteristics explain about half of the unconditional correlation. It is useful to consider this in conjunction with the corresponding change in the R^2 : In the third row, we see that it is 0.14 for the regression with all controls. This suggests that the observable characteristics explain a relatively small proportion of the overall variance in the outcome variable, but at the same time explain a large portion of the correlation between taking the lab and the outcome.

We take this as strong evidence that the unconditional correlation between taking the lab and completing the course is likely due to selection on both observable and unobservable characteristics. Because the R^2 in row 3 is 0.14, there is still much variation left in the dependent variable that could be explained with more controls. If the change in the coefficient between the unconditional and full controls scenario is any indication, we would expect the conditional correlation to keep declining if we were able to add all possible regressors to this estimation, that is, including Z_{ict} as control variables. Thus, we could consider our Full Controls estimate as an upper bound on the true causal effect. If we were able to add other characteristics to the equation, this estimate would likely shrink.

Turning to the estimates of the relationship between final grade and the lab treatment (column 2), the outcome variable is the numeric four-point version of the “A”–“F” grading scale. In the first row (and also apparent in Figure 1), we find that Lab students finish with about a third of a letter grade higher, on average, and are less likely to finish with a “D” or “F”. After conditioning for the lecture section, the correlation declines slightly to 0.32, and after conditioning on the rest of the observable characteristics, Lab students’ final grades are only about one-tenth of a letter grade higher. The conditional correlation is about one-third the size of the unconditional correlation, which suggests that two-thirds of the initial correlation (shown in row 1) is explained by the observable characteristics we incorporated in our estimate. Again, we regard this as an upper bound on the true causal effect of the lab treatment on the final grade. However, we consider the pattern between the unconditional correlation and the full controls correlation to suggest that much of the remaining conditional correlation is likely due to omitted variable bias.

The pattern holds in the final two columns as well, which depict the probability of receiving an “A” in the class and the probability of receiving an “A” or a “B” in the class. Unconditionally, Lab students are about 10 percentage points more likely to receive an “A” or “B” in the course than No Lab students, a sizable effect with practical significance if it were a causal relationship. However, after adding all controls, we find that the conditional correlation shrinks to about 3 percentage points. Again, this might be considered the upper bound on the true causal effect of taking the lab on the probability of receiving an “A” or a “B” in the course, but the pattern suggests that, if we were able to control for more variables, the size of this correlation would shrink further.

COEFFICIENT STABILITY BY STUDENT CHARACTERISTICS

Figure 2 displays the same set of coefficients from the regressions of final grade on the lab treatment for several subgroups: no controls, including classroom fixed effects, and full controls. This allows us to compare the stability of coefficients as we add controls across subgroups and to assess whether the potential for a causal relationship is stronger among certain groups. We estimate heterogeneous effects by gender, race, urbanicity, age, college GPA terciles, and Pell eligibility.

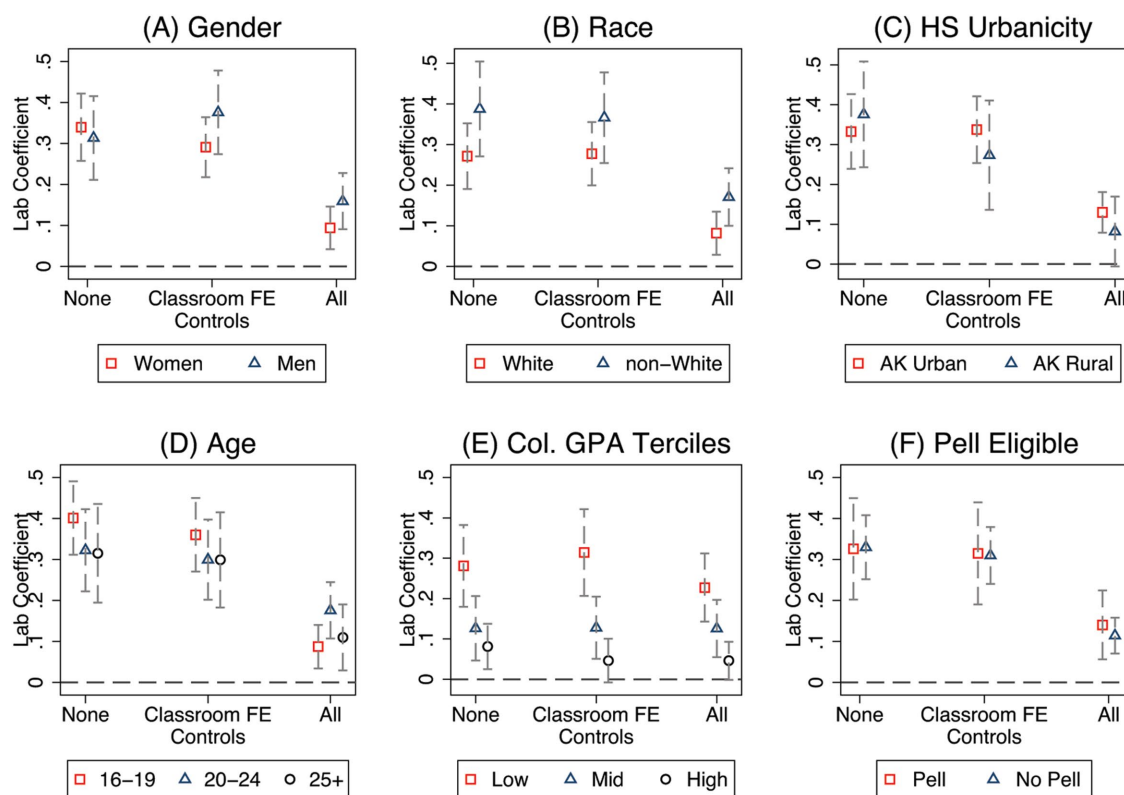


FIGURE 2. Stability of coefficients. The pattern in all six figures shows that, as controls are added to the equation, the effect of the lab reduces significantly, and this trend holds across demographic groups. FE, fixed effects.

The pattern of the correlation between the lab treatment and final grades looks similar across all subgroups. For example, Figure 2D shows the coefficients from three different groups of students: those age 16 to 19, 20 to 24, and 25 or older. All three of the first set of estimates to the left side of the panel fall between 0.315 and 0.4. Similarly, after adding all controls, the estimates all fall within the range of 0.087 to 0.18. The most stable of the three is the coefficient on students aged 20–24, which only decreases from 0.32 to 0.18 after adding all controls. The highest conditional correlation across all subgroups was for students in the lowest tercile of college GPAs; other conditional correlations that are relatively high are for the estimates among students aged 20–24, males, and non-White students. Interpreting these as upper bounds on the true causal effect of taking the lab, these estimates suggest that, if there is a causal relationship for any subgroup, it is more likely among these groups. However, even in these cases, the coefficients decline substantially from the unconditional case to the fully conditional case, suggesting that if there is a causal relationship, it is very small. On the other hand, in some other cases, particularly the estimates for urban and rural subgroups, the 95% confidence interval does not reject a null effect, and in these cases, we can essentially rule out a causal relationship.

DISCUSSION

In our literature review, we found only one relatively recent example of an experimental design that used the lab itself as a treatment to test learning outcomes; Freedman's (1997) paper on ninth-grade physics used a sample of 270 students randomly

assigned to treatment (lab experience) or control (no lab). The lab experience did not affect grades, but it did improve performance on content tests (midterms and finals) and science attitudes. Going back further, Toothacker (1983) reviewed additional studies explicitly testing the effects of labs: Kruglak (1952) found no significant difference on content knowledge tests between physics students who took labs and those who did not (though the lab group unsurprisingly scored better on lab skills). Similarly, Bradley (1965) found little difference between students randomly assigned to labs or to demonstration sections. Dubravac (1979) found no significant difference on final exams between chemistry students who had a lecture-lab experience and those who watched films during the scheduled lab time. And when Saunders and Dickinson (1979) randomly assigned first-year biology students to lecture-only, lecture-recitation, and lecture-lab groups, they found no significant difference on content knowledge tests between the recitation and lab groups; however, both scored better than the lecture-only group. Though dated, one interpretation of these findings would be that—the limitations of content knowledge tests notwithstanding—lab courses do not support content knowledge gains beyond what is gained in lecture instruction. An alternate interpretation would be that the lab's unique goals are not reflected in these assessments.

Despite the attention we gave to the variety of lab goals in the *History of Science Labs* section, the lab is still largely regarded as a strategy to reinforce content covered in lecture. As our syllabus review noted, the stated student learning outcomes at UAA are nearly identical in the lab and lecture, and the lab

activities parallel lecture content; thus, we expected our analysis to give evidence of a causal relationship, and at first glance, our findings are curious. However, science educators have problematized the curriculum's overemphasis on content knowledge outcomes: If labs primarily reinforce lecture content, they do so at the expense of metacognitive and affective outcomes (Fensham, 1984), and faculty have actively worked against this pitfall. Our analysis used course grades as the dependent variable, and as we note, the grading scheme applied focused primarily on content knowledge and low-level cognitive skills of recall/recognize with minimal application. Course grades tabulated in this manner fail to assess the critical-thinking and synthesis skills that labs seek to provide in the inquiry-based era. That students who self-select into the lab course would be able to perform equally well on low-level cognitive skills tests without the experience of a lab is not unreasonable. On the other hand, the literature documents that labs have affective outcomes and help students develop more positive attitudes about science. If the lab at UAA were achieving this, we would expect there to be a strong relationship between the lab experience and course completion outcomes (see Caine *et al.*, 2009), but our data did not reflect this either.

The finding that the lab experience is not differently impactful for specific subpopulations of students was also unexpected. Starting in the 1980s, educators noted that although Western (competitive) approaches had dominated science education (Hodson, 1993), the student body was increasingly heterogeneous, and science educators needed to serve diverse learners (Hofstein and Lazarowitz, 1986). Scholars suggested that lab activities underscored the social aspects of learning (Hodson and Hodson, 1998) through more cooperative approaches and were a way to engage diverse students (Hofstein and Lunetta, 2004)—and were thus especially effective teaching strategies for low-income and minority students (Freedman, 1997). This thinking has held in more recent scholarship in introductory biology comparing outcomes of active-learning and inquiry-based pedagogies across low-income and economically stable students (Haak *et al.*, 2011), for academically lower-performing students (Beck and Blumer, 2012), and for underrepresented minorities (URMs; Han *et al.*, 2016; Snyder *et al.*, 2016), with results noting reduced achievement gaps and greater course success rates. To the extent that labs at UAA adhere to an inquiry-based framework, we would expect the lab treatment to be especially beneficial for non-White students, but our data suggest that the lab's effect on course performance is no different for URMs or high-need students than for the rest of the student population. As increasing the participation of URMs is a significant initiative for science, technology, engineering, and mathematics education (Crisp *et al.*, 2009; President's Council, 2012), this is an interesting and unexpected finding.

LIMITATIONS

We recognize the limitation of using course completion and final grades as our outcomes; the real interest for science educators is student learning in the five dimensions we identified in our literature review. Prior research has documented that, beyond content knowledge, labs support metacognition (Karelina and Etkina, 2007; Beck and Blumer, 2016), help students think and work more independently and to do more in-depth data analyses (Howard and Miskowski, 2005), improve quanti-

tative statistical analysis skills (Goldstein and Flynn, 2011), improve scientific reasoning skills (Feyzio lu, 2009; Beck and Blumer, 2012; Brownell *et al.*, 2012), support the development of practical skills (Brownell *et al.*, 2013), and cultivate positive attitudes about science and science learning (Thompson and Soyibo, 2002; Howard and Miskowski, 2005; Beck and Blumer, 2012; Brownell *et al.*, 2013; Kloser *et al.*, 2013). While we are not conflating the final grade with student learning, our syllabus review notes that the final grades we used were primarily composed of content knowledge assessments, which is only one of many objectives that labs are meant to achieve. Our post hoc analysis of institutional data did not allow for a more accurate measure of the breadth of explicit student learning outcomes.

Additionally, beyond our syllabus review, we cannot account for how the treatment, the lab courses, were taught. Though our analysis controlled for such variables as lecture instructor quality and class size, we do not know the quality of the lab experience, or the extent to which active- or inquiry based-learning activities recommended as current best practices were applied in practice. We were unable to account for prior exposure to other labs, for example, in high school or in other college classes. If the labs are more focused on the nature of science and science literacy (with less attention to discrete concepts), it is possible that the experience of a lab associated with another science class transferred to the BIOL 102 at UAA experience. That could be another valuable conditioning variable for future analyses, and could speak to cross-disciplinary applications of science literacy. We also note that BIOL 102 is a general education course for non-science majors, and thus our results may not be transferrable to introductory science for majors, who have different interests and intent for their studies. Finally, because our analysis is limited to a single site, the results may not be generalizable or transferrable to other settings, and replication studies at other institutions with similar general education requirements are warranted.

IMPLICATIONS AND RECOMMENDATIONS

Despite noted limitations that we could identify (and likely others that readers will identify), the study's significant sample, including diversity in the data set of students and instructors, the number of variables we were able to use in conditioning, and the methodological rigor of the analysis, addresses a noted gap in the extant literature and offers some recommendations for research and practice.

Advising and Academic Supports

We expected taking a lab to correlate with improved course performance. At institutions like ours, where the labs are optional for students, our data do not suggest that labs as currently administered will help students complete the class or earn a higher grade. Thus, faculty and staff who advise students—especially students whose prior academic performance suggests that they need additional supports—should encourage those students to seek other academic support programs like tutoring or Supplemental Instruction that are specifically designed to achieve this objective (see Dawson *et al.*, 2014).

Research

More research on labs that addresses the noted limitations of our study as well as the gaps in the extant literature is

warranted. Future research should explore the role of the lab in facilitating its intended learning outcomes, including science literacy and appreciation for science; the efficacy of the lab at different levels and in different disciplines; and the effects of labs delivered using a variety of pedagogical approaches.

Science Department Pedagogy and General Education Objectives

Given the pairing and parallel nature of lecture and lab, if there are specific general education outcomes from a lab, there is an opportunity to make those goals explicit and to provide evidence that they are being achieved (Hofstein and Lunetta, 2004). Without clear objectives, studies like ours will be cited to suggest that labs are ineffective—and that would be a gross overextrapolation of our data. On the other hand, if students are able to achieve content knowledge outcomes sans a lab experience and the intent of general science education for nonmajors is to increase science literacy, our data suggest that other activities, including those that encourage students to explore concepts across disciplines, could also be a complement to the lecture. Perhaps there is a place for an experiential science lab for nonmajors, tied to a *discipline* rather than a specific lecture course, where the goals are truly independent and the general education outcome is experiencing the process of science (see also Hodson, 1992).

CONCLUSION

The history and evolution of science labs is important context as we (re)consider their contemporary purposes. The rationales for a science lab are multifaceted; these include affective gains that interest students in science as a discipline and help them to understand its nature. Though we did not test these key objectives in our study, if the lab does these things, we would expect engagement and participation in science classes to increase commensurately, and that this would be reflected—at least in part—in course completion or final grades.

There is a literature—essays and empirical studies—that questions the role of labs as a pedagogical approach, especially for nonmajors at the introductory level. Though labs are generally regarded as a “fundamental tenet of our dogma” (Thomas, 1972), in the 1970s and 1980s, some educators began to interrogate both the effectiveness and the role of laboratory work, and “the case for the laboratory was not as self-evident as it seemed” (Bates, 1978, p. 56; see also Blosser, 1988; Hodson, 1990). While some questioned their utility, others outright advocated for the elimination of labs. Criticisms included that labs are expensive to deliver in time and in dollars (Stone, 1972; Tamir, 1976; Kirschner and Meester, 1988; Reid and Shah, 2007); that labs had not empirically demonstrated their effectiveness in achieving stated outcomes (see Hofstein, 2004); and that, at the lower levels, students do not have commensurate content or background knowledge to truly gain from an inquiry-based and self-directed approach (Toothacker, 1983).

We are not advocating such a position. Though our data do pose some good questions for discussion, we share them in the spirit of what we value: our commitment to students and to science. The limitations in our data notwithstanding, our study suggests an opportunity to explore a more fundamental question: What are the science learning objectives for an undergraduate liberal arts education? Within that framework, we can

consider the role of the lab itself. We realize that the question that drove part of our literature review, What is the purpose of a lab in an undergraduate liberal arts education?, front-loaded the solution before defining the problem or need itself. With a stronger sense of its function and purpose, we encourage researchers and educators to be explicit about the goals and learning objectives for labs, to align the curricula, and to document the effectiveness of these activities through research and assessment (see also Hofstein, 2004).

As we consider the aims of the curriculum, we also implore educators to consider and underscore the civic emphasis in science education (see also Rudolph, 2014). In a sociopolitical context that characterizes scientific analysis as “fake news,” that challenges the legitimacy of such topics as climate change or human evolution, and that prefers opinion over objectivity (Marchi, 2012), the need for civic engagement with the sciences is more critical than ever. We implore educators to consider the role of the laboratory both as it relates to its component course and in the broader context of science literacy and to research its efficacy vigorously.

ACKNOWLEDGMENTS

The authors wish to thank Yuan-Fang Dong and Mist d’June Gussak for their assistance in accessing institutional data and Zeynep Kılıç and Rebeca Maseda for their thoughtful comments. This research was funded in part by National Science Foundation Improving Undergraduate Stem Education (IUSE) grant no. 1823935. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- Adami, G. (2006). A new project-based lab for undergraduate environmental and analytical chemistry. *Journal of Chemical Education*, 83(2), 253–256.
- Adams, D. L. (1998). What works in the nonmajors’ science laboratory. *Journal of College Science Teaching*, 28(2), 103–108.
- Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113(1), 151–184.
- American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC.
- Barrow, L., & Rouse, C. E. (2018). Financial incentives and educational investment: The impact of performance-based scholarships on student time use. *Education Finance and Policy*, 13(4), 419–448.
- Bates, G. R. (1978). The role of the laboratory in secondary school science programs. *What research says to the science teacher*. New York: Columbia University Teachers College.
- Beck, C. W., & Blumer, L. S. (2012). Inquiry-based ecology laboratory courses improve student confidence and scientific reasoning skills. *Ecosphere*, 3(12), 1–11.
- Beck, C. W., & Blumer, L. S. (2016). Alternative realities: Faculty and student perceptions of instructional practices in laboratory courses. *CBE—Life Sciences Education*, 15(4), 1–10.
- Beck, C., Butler, A., & da Silva, K. B. (2014). Promoting inquiry-based teaching in laboratory courses: Are we meeting the grade? *CBE—Life Sciences Education*, 13(3), 444–452.
- Blosser, P. E. (1983). What research says: The role of the laboratory in science teaching. *School Science and Mathematics*, 83(2), 165–169.
- Blosser, P. E. (1988). Labs: Are they really as valuable as teachers think they are? *Science Teacher*, 55(5), 57–59.

- Bradley, R. L. (1965). Lecture demonstration versus individual laboratory work in a general education science course. *Journal of Experimental Education*, 34(1), 33–42.
- Browne, L. M., & Blackburn, E. V. (1999). Teaching introductory organic chemistry: A problem-solving and collaborative-learning approach. *Journal of Chemical Education*, 76(8), 1104–1107.
- Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). Undergraduate biology lab courses: Comparing the impact of traditionally based “cookbook” and authentic research-based courses on student lab experiences. *Journal of College Science Teaching*, 41(4), 36–45.
- Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. J. (2013). Context matters: Volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses. *Journal of Microbiology & Biology Education*, 14(2), 176–182.
- Bybee, R. (2015). Scientific literacy. In Gunstone, R. (Ed.), *Encyclopedia of science education* (pp. 944–947). New York: Springer.
- Bybee, R. (2000). Teaching science as inquiry. In Minstrel, J., & Van Zee, E. H. (Eds.), *Inquiry into inquiry learning and teaching in science*. Washington, DC: American Association for the Advancement of Science.
- Caine, R. N., Caine, G., McClintic, C., & Klimek, K. J. (2009). *12 Brain/mind learning principles in action: Developing executive functions of the human brain*. Thousand Oaks, CA: Corwin Press.
- Carnduff, J., & Reid, N. (2003). *Enhancing undergraduate chemistry laboratories: Pre-laboratory and post-laboratory exercises*. London: Royal Society of Chemistry.
- Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, 46(4), 924–942.
- Dawson, P., van der Meer, J., Skaliky, J., & Cowley, C. (2014). On the effectiveness of supplemental instruction: A systematic review of supplemental instruction and peer-assisted study sessions literature review between 2001 and 2010. *Review of Educational Research*, 84(4), 609–639.
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37(6), 582–601.
- Dubravcic, M. F. (1979). Practical alternatives to laboratory in a basic chemistry course. *Journal of Chemical Education*, 56(4), 235–237.
- Farrell, C. R. (2016). The Anchorage mosaic: Racial and ethnic diversity in the urban North. In Barnett, J. K., & Hartman, I. C. (Eds.), *Imagining Anchorage: The making of America's Northernmost metropolis* (pp. 374–391). Anchorage, AK: University of Alaska Press.
- Fensham, P. J. (1984). Current research in chemical education. In Waddington, D. J. (Ed.), *Teaching school chemistry*. New York: UNIPUB.
- Feyzioğlu, B. (2009). An investigation of the relationship between science process skills with efficient laboratory use and science achievement in chemistry education. *Journal of Turkish Science Education*, 6(3), 114–132.
- Freedman, M. P. (1997). Relationship among laboratory instruction, attitude toward science, and achievement in science knowledge. *Journal of Research in Science Teaching*, 34(4), 343–357.
- Garnett, P. J., Garnett, P. J., & Hackling, M. W. (1995). Refocusing the chemistry lab: A case for laboratory-based investigations. *Australian Science Teachers Journal*, 41(2), 26–32.
- Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. *American Educational Research Journal*, 40(3), 731–767.
- Goldstein, J., & Flynn, D. F. (2011). Integrating active learning & quantitative skills into undergraduate introductory biology curricula. *American Biology Teacher*, 73(8), 453–461.
- Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: Gold standard, or merely standard? *Perspectives in Biology and Medicine*, 48(4), 516–534.
- Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, 332(6034), 1213–1216.
- Han, S., Capraro, R. M., & Capraro, M. M. (2016). How science, technology, engineering, and mathematics project based learning affects high-need students in the US. *Learning and Individual Differences*, 51, 157–166.
- Hart, O. M. (2018). Are lab courses worth the investment: Factors that affect student success in a biochemistry lab course. *FASEB Journal*, 32(1), 663–633.
- Hodson, D. (1988). Toward a philosophically more valid science curriculum. *Science Education*, 72(1), 19–40.
- Hodson, D. (1990). A critical look at practical work in school science. *School Science Review*, 70(256), 33–40.
- Hodson, D. (1992). In search of a meaningful relationship: An exploration of some issues relating to integration in science and science education. *International Journal of Science Education*, 14(5), 541–562.
- Hodson, D. (1993). In search of a rationale for multicultural science education. *Science Education*, 77(6), 685–711.
- Hodson, D., & Hodson, J. (1998). From constructivism to social constructivism: A Vygotskian perspective on teaching and learning science. *School Science Review*, 79(289), 33–41.
- Hoefnagels, M. (2015a). *Biology: The essentials*. New York, NY: McGraw Hill Education.
- Hoefnagels, M. (2015b). *Biology: The essentials* (2nd ed). New York, NY: McGraw Hill Education.
- Hofstein, A. (2004). The laboratory in chemistry education: Thirty years of experience with developments, implementation, and research. *Chemistry Education Research and Practice*, 5(3), 247–264.
- Hofstein, A., & Lazarowitz, R. (1986). A comparison of the actual and preferred classroom learning environment in biology and chemistry as perceived by high school students. *Journal of Research in Science Teaching*, 23(3), 189–199.
- Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of Educational Research*, 52(2), 201–217.
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88(1), 28–54.
- Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: The state of the art. *Chemistry Education Research and Practice*, 8(2), 105–107.
- Holbrook, J., & Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental and Science Education*, 4(3), 275–288.
- Howard, D. R., & Miskowski, J. A. (2005). Using a module-based laboratory to incorporate inquiry into a large cell biology course. *Cell Biology Education*, 4(3), 249–260.
- Karelina, A., & Etkina, E. (2007). Acting like a physicist: Student approach study to experimental design. *Physics Education Research*, 3(2), 1–12.
- Kirschner, P. A., & Meester, M. A. M. (1988). The laboratory in higher science education: Problems, premises and objectives. *Higher Education*, 17(1), 81–98.
- Klainin, S. (1988). Practical work and science education. In Fensham, P. (Ed.), *Development and dilemmas in science education* (pp. 169–188). Bristol, PA: Taylor & Francis.
- Kloser, M. J., Brownell, S. E., Shavelson, R. J., & Fukami, T. (2013). Effects of a research-based ecology lab course: A study of nonvolunteer achievement, self-confidence, and perception of lab course purpose. *Journal of College Science Teaching*, 42(3), 72–81.
- Krajcik, J., Mamlok, R., & Hug, B. (2001). Modern content and the enterprise of science: Science education in the twentieth century. In Corno, L. (Ed.), *Education across a century: The centennial volume* (pp. 205–238). Chicago: University of Chicago Press.
- Kruglak, H. (1952). Achievement of physics students with and without laboratory work. *American Journal of Physics*, 21(1), 14–16.
- Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. In Abell, S., & Lederman, N. (Eds.), *Handbook of research on science education* (pp. 393–441). Mahwah, NJ: Erlbaum.
- Mader, M., & Windelspecht, M. (2014). *Essentials of biology* (4rd ed). New York, NY: McGraw Hill Education.
- Mader, M., & Windelspecht, M. (2017a). *Essentials of biology* (3rd ed). New York, NY: McGraw Hill Education.
- Mader, M., & Windelspecht, M. (2017b). *Essentials of biology* (5rd ed). New York, NY: McGraw Hill Education.

- Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity." *Journal of Communication Inquiry*, 36(3), 246–262.
- Matz, R. L., Rothman, E. D., Krajcik, J. S., & Banaszak Holl, M. M. (2012). Concurrent enrollment in lecture and laboratory enhances student performance and retention. *Journal of Research in Science Teaching*, 49(5), 659–682.
- McQueen, J. A. (2017). *The effects of biology lab delivery mode on academic achievement in college biology* (Doctoral dissertation). Retrieved from ProQuest. 10259993.
- Meyers, C., & Brandt, W. C. (2014). *Implementation fidelity in education research: Designer and evaluator considerations*. New York: Routledge.
- Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE—Life Sciences Education*, 9(4), 435–440.
- National Research Council. (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. Washington, DC: National Academies Press.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187–204.
- President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering and mathematics*. Washington, DC: U.S. Government Office of Science and Technology.
- Puttick, G., Drayton, B., & Cohen, E. (2015). A study of the literature on lab-based instruction in biology. *American Biology Teacher*, 77(1), 12–18.
- Reid, N., & Shah, I. (2007). The role of laboratory work in university chemistry. *Chemistry Education Research and Practice*, 8(2), 172–185.
- Rudolph, J. L. (2014). Dewey's "science as method" a century later: Reviving science education for civic ends. *American Educational Research Journal*, 51(6), 1056–1083.
- Saunders, W. L., & Dickinson, D. H. (1979). A comparison of community college students' achievement and attitude changes in a lecture-only and lecture-laboratory approach to general education biological science courses. *Journal of Research in Science Teaching*, 16(5), 459–464.
- Snyder, J. J., Sloane, J. D., Dunk, R. D., & Wiles, J. R. (2016). Peer-led team learning helps minority students succeed. *PLoS Biology*, 14(3), e1002398.
- Son, J. Y., Narguizian, P., Beltz, D., & Desharnais, R. A. (2016). Comparing physical, virtual, and hybrid flipped labs for general education biology. *Online Learning*, 20(3), 228–243.
- Stone, T. J. (1972). Endpoint: Down with labs. *Education in Chemistry*, 9(5), 204.
- Sullivan, G. M. (2011). Getting off the "gold standard": Randomized controlled trials and education research. *Journal of Graduate Medical Education*, 3(3), 285–289.
- Tamir, P. (1976). *The role of the laboratory in science teaching (Technical Report No. 10)*. Iowa City: University of Iowa Science Education Center.
- Thomas, K. W. (1972). The merits of continuous assessment and formal examinations in practical work. *Journal of Biological Education*, 6(5), 314–318.
- Thompson, J., & Soyibo, K. (2002). Effects of lecture, teacher demonstrations, discussion and practical work on 10th graders' attitudes to chemistry and understanding of electrolysis. *Research in Science & Technological Education*, 20(1), 25–37.
- Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. *School Science and Mathematics*, 90(5), 403–418.
- Toothacker, W. S. (1983). A critical look at introductory laboratory instruction. *American Journal of Physics*, 51(6), 516–520.
- University of Alaska Anchorage Office of Institutional Research. (2015). *2014–15 fact book*. Anchorage: University of Alaska Anchorage. Retrieved January 17, 2019, from www.uaa.alaska.edu/academics/institutional-effectiveness/departments/institutional-research/_documents/2015-students-final.pdf
- University of Alaska Statewide Office of Institutional Research. (2016, January). *Performance update: Developmental education at UAA*. Fairbanks, AK: Institutional Effectiveness, Engagement, and Academic Support.