In response to calls for reform in undergraduate biology education, we conducted research examining how varying active-learning strategies impacted students’ conceptual understanding, attitudes, and motivation in two sections of a large-lecture introductory cell and molecular biology course. Using a quasi-experimental design, we collected quantitative data to compare participants’ conceptual understanding, attitudes, and motivation in the biological sciences across two contexts that employed different active-learning strategies and that were facilitated by unique instructors. Students participated in either graphic organizer/worksheet activities or clicker-based case studies. After controlling for demographic and presemester affective differences, we found that students in both active-learning environments displayed similar and significant learning gains. In terms of attitudinal and motivational data, significant differences were observed for two attitudinal measures. Specifically, those students who had participated in graphic organizer/worksheet activities demonstrated more expert-like attitudes related to their enjoyment of biology and ability to make real-world connections. However, all motivational and most attitudinal data were not significantly different between the students in the two learning environments. These data reinforce the notion that active learning is associated with conceptual change and suggests that more research is needed to examine the differential effects of varying active-learning strategies on students’ attitudes and motivation in the domain.

INTRODUCTION
Analyses by the President’s Council of Advisors on Science and Technology (PCAST) predict a national shortage of one million science, technology, engineering, and mathematics (STEM) professionals by 2022 (PCAST, 2012). In response to this concern, experts recommend implementing active-learning strategies (ALS) and attending to students’ noncognitive attributes (e.g., attitudes and motivation) as viable methods for increasing student retention in the STEM disciplines (Pajares and Schunk, 2001; American Association for the Advancement of Science, 2010; PCAST, 2012; Watkins and Mazur, 2013). Freeman et al.’s (2014) seminal meta-analysis provided robust support for the use of ALS, compared with traditional lecture, as a means to improve undergraduate STEM students’ academic performance. A variety of ALS exist, and their use varies by instructor. What is less clear, however, is which ALS (e.g., students formulating their own questions following a reading assignment, students participating in peer discussion, or students working collaboratively/individually on complex problems) best promote student learning and attend to...
students’ affect (i.e., learning attitudes and motivation) in the domain (Pajares and Schunk, 2001).

Interestingly, research suggests that not all instructors experience expected gains in student learning when implementing ALS (Andrews et al., 2011). For instance, Andrews et al. (2011) reported on differences in learning gains from a random sample of collegiate biology instructors, both those active in education research and content biologists, who self-reported using ALS. After accounting for multiple variables (e.g., student-rated course difficulty, instructor’s years of experience, and class size), the data demonstrated that students whose instructors were content biologists displayed significantly fewer learning gains than those students enrolled in courses facilitated by instructors trained in conducting educational research. The authors concluded that the biology content specialist (i.e., an individual with less pedagogical content knowledge—knowledge and skills related to how to promote student learning within a content domain; Shulman, 1986) did not understand the constructivist elements deeply enough to produce the type of learning gains typically associated with active learning. After commending Andrews et al. (2011) for their use of linear regression to reduce confounding due to instructors’ experience and the classroom environment, Theobald and Freeman (2014) noted that the study by Andrews and colleagues did not take into account student characteristics (e.g., gender and first language) that may have directly impacted their results.

To maximize instructors’ efforts and student learning, more research is needed on: 1) which ALS work for whom (e.g., students with minimal precollegiate STEM course work vs. highly prepared students) and in which settings (e.g., lower- vs. upper-division courses, small vs. large classrooms); 2) what underlying mechanisms are responsible for the cognitive gains associated with active learning (e.g., increased metacognition); and 3) how to effectively train collegiate instructors to successfully implement ALS. The question is no longer whether expository, lecture-style teaching should serve as the primary mode of instruction, but rather what type(s) of active learning are most effective in promoting student learning and affect in the domain. Specifically, our research was guided by the following goals: 1) to examine how instructors with different educational backgrounds and pedagogical training implement ALS in a large-enrollment introductory cellular and molecular biology course; and 2) to quantify the differential impacts of those distinct active-learning contexts on students’ conceptual understanding, attitudes, and motivation toward learning biology. To achieve these goals, we compared student outcomes from two biology instructors who had different educational backgrounds and who used different ALS (see Table 1).

The instructors in our study primarily used either graphic organizers/worksheets or clicker-based case studies as a means to engage students in the learning process. These student-centered instructional approaches are commonly used by both secondary and postsecondary STEM educators (e.g., Pinto and Zeitz, 1997; Edmondson, 2000; Kinchin, 2000; Weiss and Levinson, 2000; Bonney, 2015). Graphic organizers are two-dimensional representations of knowledge that can be teacher- and/or student-generated. Various forms of graphic organizers exist: Venn diagrams, flowcharts, and timelines. Students play an active role in determining which information should be included on their graphic organizers regardless of whether the structure is teacher or student-created. Instructors can use graphic organizers in a variety of ways: to provide an overview of the material to be learned, to provide a framework for new vocabulary, to provide reading cues, and/or to provide a concise review guide (Hawk, 1986). Proponents of graphic organizers argue that the use of this instructional method is correlated with a deep learning approach (Laitig, 2006) and, therefore, the promotion of meaningful learning (Novak and Gowin, 1984; Novak, 1991a,b; Watson, 1989; Okebukola, 1990). The literature on the use of graphic organizers in post-secondary contexts suggests that students who participate in activities of this nature display improved higher-order thinking, ability to determine hierarchical relationships, reading comprehension, problem-solving skills, essay writing skills, and conceptual understanding of content material (Alvermann, 1981; Kiewra et al., 1988, 1999; Robinson and Kiewra, 1995; Robinson and Schraw, 1994; Katayama and Robinson, 2000; Novak and Musonda, 1991).

Case study pedagogy, an alternate form of ALS, has likewise gained significant traction among collegiate STEM instructors, most likely due to the ease with which it can be applied in both small- and large-lecture settings (Merseth, 1991; Knechel, 1992; Herreid, 1994, 2006; Clift and Wright, 1996; Dori and Herscovitz, 1999; Flynn and Klein, 2001; Tomey, 2003; Mayo, 2004; Olgun and Adali, 2008; Wolter et al., 2011; Murray-Nsuala, 2012; Yalcinkaya et al., 2012). The formation of the University of Buffalo’s National Center for Case Study Teaching (NCCST) has provided collegiate instructors with access to more than 500 peer-reviewed case studies designed to assist faculty in incorporating problem-based learning, teaching content, and promoting critical thinking—all while using relevant and authentic science. The use of case study teaching has been shown to be an effective means to improve students’ academic performance and affect in the domain (Cliff and Wright, 1996; Bonney, 2015).

It is important to note that a variety of case study teaching methods exist (e.g., analysis case, decision case, directed case, flipped case, discussion case; see NCCST, 2016). The “clicker case,” or “interrupted case,” is the most popular type of case study teaching among collegiate science instructors (Yadav et al., 2006), most likely due to its ease of use and success in large-lecture settings (Michaelsen et al., 2002; Herreid, 2006). Instructors incorporating case studies into their classrooms present information in a narrative format followed by a series of questions (Herreid et al., 2011). These instructional methods allow students to problem solve, think like scientists, and analyze data (Herreid, 2006; Herreid et al., 2011). In terms of cognitive skills, case study teaching has been shown to improve students’ analytical skills, higher-order thinking, and exam performance (Herreid, 1994; Herreid et al., 2011). Additionally,
students participating in a case-based, active-learning environment show increased motivation and engagement (Bonney, 2015)—key factors linked to academic performance (Pintrich and Schunk, 2012; Yalcinkaya et al., 2012).

Equally important to exploring the impact of ALS on students’ conceptual understanding is discovering the impact of varying instructional strategies on students’ affect in the domain. Students’ attitudes and motivation toward a subject have been shown to impact their persistence within a STEM major (e.g., Adams et al., 2006). In addition to positive attitudes, motivated students are more likely to display behaviors (e.g., attending class, asking questions, seeking advice, participating in study groups) that increase their probability of academic success (Pajares, 1996, 2002; Pajares and Schunk, 2001). Therefore, it is important when comparing the differential effects of various ALS employed by instructors to not only examine their impact on students’ conceptual understanding but to also explore their influence on noncognitive student outcomes (e.g., attitudes and motivation).

A great deal of research has been conducted that demonstrates the benefits of ALS as compared with lecture (Crouch and Mazur, 2001; Freeman et al., 2014). However, few studies have focused on comparing the differential effects of varied forms of ALS. Expanding upon previous work, this study takes into account conceptual understanding and noncognitive factors when examining relative effectiveness of implementing different ALS in an introductory, undergraduate cell and molecular biology course. Specifically, we sought to determine: 1) the extent to which two different professors are using ALS in their classrooms; and 2) the extent to which the varied ALS implemented by these instructors (e.g., worksheets/graphic organizers vs. clicker-based case studies) differentially impact students’ learning, attitudes, and motivation in an introductory cell and molecular biology course.

**METHODS**

**Context**

This research was conducted at a midsize university in the West. Data were collected from all (2) instructors who taught the same introductory cell and molecular biology course. The major topics covered in this course included 1) principles of chemistry; 2) cells, energy, and metabolism; 3) replication, transcription, translation, and gene expression; and 4) genetics. The course is required for all biology, chemistry, allied health, and nutrition majors and also serves as a liberal arts core for non-STEM majors. Class sizes ranged between 100 and 250 students.

**Instructor Participants**

The two instructors (pseudonyms: Jennifer and Matthew) recruited to participate in the study were first asked to provide basic demographic information about their educational backgrounds, teaching experience, and the types of ALS they typically used in their classrooms (Table 1). The ALS used were confirmed through observation of the course (see Data Collection and Analysis). Though the instructors differed in the total years they had been teaching in the biological sciences, both faculty had taught introductory cell and molecular biology at the university for at least three semesters. Matthew and Jennifer both pursued bachelor’s and master’s degrees in biology; however, Matthew’s graduate work was in science education, whereas Jennifer’s was in biology. In addition, both faculty made use of a variety of ALS, used clickers during standard lectures, used the same textbook and similar course notes, and shared a common syllabus.

**Student Participants**

Student participants ($n_{	ext{total}} = 132$) were selected through matched comparisons from a convenience sample consisting of all students enrolled in either Jennifer’s or Matthew’s course ($n_{	ext{Jennifer}} = 66$; $n_{	ext{Matthew}} = 66$). In an effort to account for potential bias introduced as a result of heterogeneity between section cohorts, participants were matched on several demographic and psychosocial factors (e.g., index score [a measure of college readiness], race/ethnicity, gender, major, presemester self-efficacy [a measure of belief in one’s ability]) previously identified as predictors of student success in the STEM domains (Tai et al., 2006). Demographic data were obtained through the center for institutional reporting at the site at which this research occurred. Presemester self-efficacy data were determined from the Science Motivation Questionnaire II—Biology, administered as described below (Glynn et al., 2011; see Attitudes and Motivation subsection under Data Collection and Analysis). In addition, only those students enrolled in the course for the first time and who completed all aspects of the data-collection protocol were included in our analyses. The former step was implemented intentionally as a mechanism to control for participants’ prior exposure to course content.

Specifically, participants in Jennifer’s section were first matched to one or more participants in Matthew’s section on those demographic and psychosocial variables referenced above (Table 2). This resulted in direct matching on all variables excluding index score, self-efficacy, and self-determination (a measure of one’s ability to persist in a given task). To account for variation in these factors, we retained participants in Matthew’s section as potential matches only if their scores on the aforementioned constructs were within one-half SD of those scores reported for the proposed aligned matches within Jennifer’s section. In instances in which more than one suitable match within Matthew’s section was identified, a random number generator was used to create a one-to-one pairing. To determine whether a significant difference in index score, self-efficacy, and/or self-determination existed between groups subsequent to the matching procedure, we performed a series of independent $t$ tests. These data revealed no statistically significant, between-cohort differences on any of the above variables ($p > 0.395$ for all comparisons).

**Instruments**

**Classroom Observation Protocol for Undergraduate STEM.** The Classroom Observation Protocol for Undergraduate STEM (COPUS) was used to measure differences in instructors’ pedagogical strategies (Smith et al., 2013). The COPUS is a qualitative, periodic-interval instrument designed to create a profile of instructor and student practices in collegiate science classrooms and to provide a means for quantifying instructor and student behaviors. We used the COPUS for the latter purpose. With the COPUS, researchers measure the frequency of 12 possible instructor behaviors and 13 possible student behaviors during a class period (e.g., students discussing clicker questions in groups of two or more, instructor moving through class guiding...
ongoing student work). In each 2-minute interval during the class, every behavior (instructor or student) that is observed is recorded. Instructor and student behaviors (COPUS codes) are reported as a percentage of time engaged in an activity during class (Lund et al., 2015). Based on the purpose of our study, one modification was made. In addition to recording student group activities (discuss clicker questions in groups, work on worksheet groups, and other group activities), we also calculated the total percentage of time students engaged in group work by combining those three categories.

Learning Gains. The Introductory Molecular and Cell Assessment (IMCA) was administered in a pre–post manner to allow us to measure student learning gains (Hake, 1998) over the course of the semester. The IMCA consists of 24 multiple-choice questions administered over one 30-minute period, with items covering the breadth of content traditionally encountered in an introductory biology survey course (Shi et al., 2010). Students completed the IMCA in its entirety; 18 questions were used for assessment purposes, and the remaining six questions, which were not directly addressed during the course, served as a measure of internal validity (Cronbach’s α = 0.752).

Attitudes and Motivation. By administering the Colorado Learning Attitudes in Science Survey–Biology (CLASS-Bio; Semsar et al., 2011) and the Science Motivation Questionnaire II–Biology (BMQ; Glynn et al., 2011), we were able to quantitatively measure changes in students’ attitudes and motivation, respectively. The CLASS-Bio and BMQ have been used previously in both traditional and active learning–based environments to examine shifts in the aforementioned constructs. Both instruments have been demonstrated to be valid in populations similar to ours (Cronbach’s α = 0.855 and 0.844, respectively, within the research context described herein).

The CLASS-Bio consists of 31 Likert-item questions designed to examine the extent to which students agree with expert responses on seven scales: Real-World Connections, Problem-Solving Difficulty, Enjoyment, Problem-Solving Effort, Conceptual Connections, Problem-Solving Strategies, and Reasoning. A shift in percent favorable scores is reported. To determine a shift in percent favorable responses, a pre- and postassessment must be administered and percent favorable scores compared between the time periods. An individual’s percent favorable score represents the proportion of responses provided by the student that align with those provided by experts in the field (i.e., someone holding a PhD in biology). An increase in percent favorable responses (i.e., student responses that approximate expert-like responses) over time is indicated by a positive shift, whereas a decrease in percent favorable responses represents a decrease in expert-like thinking. Previous research in introductory physics, chemistry, and biology education has demonstrated that undergraduate students receiving traditional or active learning–style instruction often display negative shifts in their attitudes toward STEM (Redish et al., 1998; Perkins et al., 2005; Wieman, 2007; Semsar et al., 2011). Positive shifts in students’ attitudes have been observed in classes implementing pedagogical techniques aimed at addressing epistemological issues (e.g., see Hammer 1994; Redish et al., 1998; Perkins et al., 2005; Otero and Gray, 2008).

In comparison, the BMQ consists of 25 Likert-item questions regarding intrinsic and extrinsic factors related to students’ motivation in the biological sciences (career and grade motivation, self-efficacy, self-determination, intrinsic motivation). A mean score is reported for each motivational category assessed. Higher scores represent a higher level of motivation within the specified category. The creators of the BMQ validated the instrument for use in majors and nonmajors biology contexts and suggested it be used as tool to measure both differences in motivations between populations and longitudinal changes in motivation (Glynn et al., 2011). The BMQ has primarily been used to determine whether different motivational levels exist between populations (e.g., Campos-Sánchez et al., 2014). Therefore, the current literature base does not indicate whether it is common for students to display an increase or decrease in motivation while completing an introductory biology course. Both diagnostics can be completed in one 45-minute time period.

Data Collection and Analysis

Pedagogical Style. With their consent, instructors were video-recorded nine times throughout the duration of the semester, with course topics being identical between instructors during the episodes in which data were collected. Two researchers independently coded the videos using the COPUS
TABLE 3. Comparison of instructor behaviors

<table>
<thead>
<tr>
<th>Category</th>
<th>Jennifer (median)</th>
<th>Matthew (median)</th>
<th>Mann-Whitney U</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student group work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicker question discussions</td>
<td>35%</td>
<td>4%</td>
<td>2.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Group worksheet exercises</td>
<td>0%</td>
<td>24%</td>
<td>13.50</td>
<td>0.014</td>
</tr>
<tr>
<td>Other group activities</td>
<td>0%</td>
<td>0%</td>
<td>26.00</td>
<td>0.222</td>
</tr>
<tr>
<td>Total group work</td>
<td>38%</td>
<td>27%</td>
<td>33.00</td>
<td>0.546</td>
</tr>
<tr>
<td>Instructor behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real-time writing on the board</td>
<td>0%</td>
<td>32%</td>
<td>4.50</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Posing questions (non-clicker)</td>
<td>76%</td>
<td>31%</td>
<td>10.00</td>
<td>0.006</td>
</tr>
<tr>
<td>Following up on group work</td>
<td>31%</td>
<td>15%</td>
<td>17.00</td>
<td>0.040†</td>
</tr>
<tr>
<td>Listening to and answering student questions</td>
<td>13%</td>
<td>4%</td>
<td>17.00</td>
<td>0.040†</td>
</tr>
<tr>
<td>Engaged in lecturing</td>
<td>90%</td>
<td>88%</td>
<td>36.00</td>
<td>0.730</td>
</tr>
<tr>
<td>Moving around the room and guiding student work</td>
<td>15%</td>
<td>20%</td>
<td>38.00</td>
<td>0.863</td>
</tr>
<tr>
<td>Working one-on-one with students</td>
<td>20%</td>
<td>31%</td>
<td>34.50</td>
<td>0.605</td>
</tr>
<tr>
<td>Waiting</td>
<td>9%</td>
<td>0%</td>
<td>22.00</td>
<td>0.113</td>
</tr>
</tbody>
</table>

*Values represent the median of the percentage of time spent on various activities over nine class periods. They do not add up to 100%, because multiple behaviors can be observed during any 2-minute interval.

†Not significant following Bonferroni correction (p = 0.006).

(Smith et al., 2013). High interrater reliability was observed (κ = 0.944; Landis and Koch, 1977). Any instructor codes that were not used by at least one of the instructors in more than half of the classes or that took up less than 5% of class time were removed from the statistical analysis (e.g., showing or conducting a demonstration, administration, and other). Of the remaining teacher codes, “teacher poses clicker question” was also removed, because it was directly related to the “students discuss clicker questions in a group” code. In Jennifer’s and Matthew’s classrooms, clicker questions were always followed up with group work (e.g., student discussion). To determine whether the instructor behaviors differed, we performed a Mann-Whitney U-test on the instructor codes. To determine whether there was a difference between the types of student group work between instructors, we compared the three student group-work codes (discuss clicker questions in groups, work on worksheets in groups, and other group activities) using a Mann-Whitney U-test. For both instructor and student behavior comparisons, alpha was set a priori to 0.050, and a Bonferroni correction was made for multiple related tests. There are eight instructor behavior categories (see Table 3, Bonferroni correction: p = 0.006) and three student group work categories (see Table 3, Bonferroni correction: p = 0.017). All statistics were run in SPSS (version 22; IBM, Armonk, NY).

Learning Gains. The IMCA (Shi et al., 2010) was administered during the first and 12th weeks of the semester during part of a lecture period. Student responses were recorded on Scantron forms, scored electronically, and entered directly into SPSS, version 22. Results were reported as a percentage score (out of the 18 concepts covered in the course). Learning gains were calculated as follows: \( g = 100 \times \text{(posttest score – pretest score)} / (100 – \text{pretest score}) \) (Hake, 1998). Independent and paired t tests were performed to determine whether a significant difference in learning gains existed between or within groups, respectively. Statistical effect sizes were calculated using Cohen’s d (Cohen, 1992), where a Cohen’s d value greater than 0.8 is considered a large effect size.

Attitudes and Motivation. The CLASS-Bio and BMQ were administered during the first and 12th weeks of the semester during part of the laboratory period associated with the course. Student responses were recorded on Scantron forms, scored electronically, and entered directly into SPSS, version 22. Using students’ IMCA scores as a covariate, we performed a multivariate analysis of covariance (MANCOVA) to compare students’ shifts in attitudes and motivation between classes.

Real-World Connections. After reviewing our attitudinal data, specifically the Real-World Connections and Enjoyment categories of the CLASS-Bio, and finding that Matthew’s class displayed more expert-like thinking compared with Jennifer’s class (see Figure 1), we elected to conduct a posteriori (previously not planned) analyses to determine whether instructors differed in their use of real-world examples or analogies.

The same nine videos used for the COPUS were also used for the a posteriori analyses. After watching recorded lectures, the coding team developed a framework for recording the instructors’ use of real-world connections (i.e., connecting biological concepts to real-world phenomena; a category within the CLASS-Bio index). This coding scheme included recording the number of times each instructor used an analogy and/or real-world example. For this coding, an analogy was defined as connecting a real-world object to a biological structure, function, or topic. For example, noting that the structure of grana in a chloroplast looks like a stack of pancakes or relating cell structure and function to a city (e.g., nucleus directs traffic and membrane-bound organelles are the workers). A real-world example was an everyday connection with the use of facts, concepts, or ideas. For example, the oxygen in our atmosphere is from the blue-green algae’s photosynthetic properties (photosynthesis lecture) or chemicals in cigarettes can cause mutations (genetics lecture).
All real-world examples and analogies made during each video-recorded class for each instructor were coded and recorded. The total number of analogies and real-world examples were recorded and reported per 50-minute class period. Instructors’ use of analogies and real-world examples were compared using a Mann-Whitney U-test with Bonferroni correction for two comparisons (p = 0.025).

**RESULTS**

**Identification of Similarities and Differences in Instructors’ Pedagogical Approaches**

In an effort to create a comprehensive and nuanced representation of each instructor’s teaching style, we generated COPUS data both to describe those activities in which students were engaged and to detail the pedagogical approaches and strategies used by the instructors during recorded sessions. These data indicated that, while the median amount of time Matthew and Jennifer implemented group work in their sections did not differ statistically, their approach to student group work did (Table 3). Jennifer regularly used clicker-based case study exercises (Herreid, 2006; NCCST, 2016). After posing each clicker question within her case studies, Jennifer provided her students with the opportunity to discuss the question with their peers (Table 3; the median time spent by Jennifer’s students on clicker question discussion was 35%, whereas the median time spent by Matthew’s students was 4%; U = 2.00, p < 0.001). In contrast, Matthew preferred to make use of a variety of worksheet-based activities to reinforce course content (examples can be found in the Supplemental Material). Students in Matthew’s class participated in worksheet-based activities a median of 24% of the time, while Jennifer’s students spent a median of 0% of the time on worksheet-based activities (Table 3). As a follow-up to his students’ group work, Matthew wrote on the document camera, particularly as a mechanism to review the material. Matthew wrote on the document camera for a median of 32% of the time, while Jennifer was not observed using the document camera (0%; Table 3; U = 4.50, p < 0.001). The difference in instructors’ use of alternate group-work approaches (e.g., one-minute papers, in-class discussions) was not found to be statistically significant (U = 33.00; p = 0.546).

Outside of student group work, Jennifer was noted to pose more non-clicker questions (i.e., call and response) to students in her class than Matthew (U = 10.00; p = 0.006; Table 3). With the exceptions of Matthew’s use of the document camera (i.e., Real-Time Writing on the Board) and Jennifer’s use of non–clicker questions, results indicate that the median percentage of time Matthew and Jennifer spent on other instructional behaviors (i.e., non–student group work) was not significantly different (Bonferroni correction of p = 0.006 for the nine non–group work behaviors; see Table 3). This included behaviors related to active learning (e.g., following up on student group work; U = 17.00; p = 0.040), answering student questions (U = 17.00; p = 0.040), moving and guiding during student work (U = 38.00; p = 0.863), one-on-one discussion with students during group work (U = 34.50; p = 0.605), and engaging in lecturing (U = 36.00; p = 0.730).

**Similar Gains in Student Learning Observed Despite Differences in Instructor Expertise and Adoption of Pedagogical Strategies**

For determination of the impact of instructor expertise and adoption of pedagogical strategies on students’ conceptual understanding in an introductory cell and molecular biology course, normalized scores on the IMCA were first computed and an independent t test subsequently performed to assess for between-group differences on the concept inventory. This analysis revealed no statistically significant difference between the student learning gains observed in Jennifer’s section of the course (M = 16.07%; SEM = 3.73%) and those observed in Matthew’s section of the course (M = 7.11%; SEM = 3.38%) (t(130) = −1.78; p = 0.077). Importantly, however, within-group analyses demonstrated that students in both sections of the course exhibited significant gains in conceptual understanding over the course of the semester (Table 4).

![FIGURE 1. Comparison of shifts in students’ attitudes stratified by instructor; *p < 0.05.](image-url)
Some Student Differences in Attitudes Were Observed in Varying Active-Learning Environments

Student responses on the CLASS-Bio were analyzed using a MANCOVA procedure to assess for pre- and postsemester shifts in participants’ attitudes in the biological sciences. When interpreting these results, it is important to note that most students, regardless of the pedagogical strategy used by an instructor (i.e., traditional lecture or ALS), display negative shifts in their STEM-related beliefs (Perkins et al., 2005; Adams et al., 2006). When students’ IMCA scores were used as a covariate, results demonstrated a statistically significant, between-group difference on two of the seven scales (Real-World Connections and Enjoyment; see Figure 1). Students in both sections displayed expected negative shifts in Real-World Connections (e.g., Semsar et al., 2011); however, students in Matthew’s section experienced a significantly less negative shift in attitudes on this factor as compared with students in Jennifer’s class. In the category of Enjoyment, Matthew’s students displayed a shift toward more expert-like attitudes, while students in Jennifer’s class displayed a shift toward more novice-like attitudes. Overall, for the categories of Real-World Connections and Enjoyment, students in Matthew’s section exhibited attitudes that more closely resemble those of experts when compared with students in Jennifer’s class. Considering that previous studies have only observed positive shifts in students’ thinking when instructors intentionally designed their curriculum to address students’ epistemology (e.g., having students reflect upon and critique their own thinking; Hammer, 1994), Matthew’s students’ shift toward more expert-like thinking in the category of Enjoyment is important to note.

To ascertain why the Real-World Connections data were different between instructors, we performed a posteriori analyses in which the total number of real-world examples and analogies used by each instructor over the course of the semester was tabulated from collected video data. These analyses indicated that, while Matthew used a significantly greater number of analogies than Jennifer (U = 12.50; p = 0.011; Table 5), no statistically significant difference in number of real-world examples (e.g., everyday connections to biology concepts) was observed between instructors after a Bonferroni correction of p = 0.025 (U = 16.00; p = 0.031; Table 5). Owing to the a posteriori design, a causal relationship cannot be confirmed between an instructor’s use of analogies in class and students’ attitudes toward Real-World Connections.

Participation in Diverse Active-Learning Exercises Results in Parallel Shifts in Student Motivation

In addition to exploring the impact of diverse instructional experiences on students’ conceptual understanding and attitudes in the domain, we further sought to examine the degree to which these varied approaches influenced students’ motivation in the biological sciences. Negative shifts, as expected based on previous research, were observed in all categories (Ding and Mollohan, 2015). When participants’ IMCA scores were used as a covariate, results from a Mann-Whitney U-test indicated no significant between-group differences on any of the five factors found on the BMQ (Figure 2).

DISCUSSION

The research presented herein contributes to the increasing amount of evidence demonstrating that students enrolled in courses that incorporate ALS display positive changes in their conceptual understanding (Freeman et al., 2014). In addition, and in alignment with previous literature, students in our study did not, on the whole, exhibit positive shifts in their attitudes and motivation within the domain (Redish et al., 1998; Perkins et al., 2005; Adams et al., 2006; Wieman, 2007; Semsar et al., 2011). Uniquely, this study demonstrates that: 1) implementation of different ALS for similar amounts of time, but in different frequencies (i.e., entire class periods once a week as opposed to approximately one-third of the class per class period), can result in similar gains in conceptual knowledge; 2) instructors with different educational backgrounds and training can effectively implement ALS; and 3) implementation of different ALS can have variable impacts on certain areas of students’ attitudes toward biology.

In terms of conceptual understanding, students in both Jennifer’s and Matthew’s classes displayed significant gains in conceptual knowledge with a large effect size (d = 0.43–0.85; see Table 4) observed for the IMCA scores (large effect size is ≥0.80; see Cohen, 1992). This is larger than found in Van Dusen et al. (2015), which averaged d = 0.306 in courses using learning assistants. However, the improvement in student learning identified within our own context is lower than that reported by both the original developers of the IMCA (Shi et al., 2010) and two other studies that used the IMCA in introductory-level biology courses (Jensen et al., 2013;
Wolkow et al., 2014). This may be because the student population at the university at which this research was conducted differs, demographically, from participants in those studies or because there was less emphasis on IMCA-related material in the context described herein (30.8% IMCA content covered in both Jennifer’s and Matthew’s sections).

We acknowledge that this study is quasi-experimental in nature, and we did not dictate which ALS Matthew and Jennifer used nor at what frequency; therefore, causality of these findings must be interpreted with caution. Matthew and Jennifer employed approximately the same amount of active learning in a large introductory biology lecture, with both using about one-third of the class period, on average, for active learning, and lecturing as the next most-common mode of instruction. However, the strategies for active learning used differed. Matthew primarily used graphic organizers and worksheets, while Jennifer used clicker-based case studies. Additionally, the way the instructors used the third of their time in ALS was different. The majority of Jennifer’s active-learning time was spent by using an entire class period (50 minutes) once a week to complete clicker-based case studies, while Matthew used active learning distributed throughout almost every class period. We recommend that future researchers design a controlled experiment to determine how the distribution of time spent on various ASL impacts students’ learning and attitudes. For example, it would be beneficial to examine students’ conceptual learning, attitudes, and motivation when the same ALS was implemented at different frequencies and durations (e.g., 15 minutes per class period three times a week compared with 45 minutes in one class period).

Furthermore, Matthew and Jennifer differ in their doctoral training and experience in teaching (Table 1). Despite these differences, it is interesting to note that there was no significant difference in conceptual learning or motivation between their students. Additionally, there was little significant difference between their students in relationship to participants’ attitudes toward biology. As described earlier, Matthew’s and Jennifer’s predominant choice of ALS (i.e., graphic organizers vs. clicker-based case studies) and the frequency and duration with which they implemented ALS varied. They also differed in other teaching behaviors, some of which relate to their different ALS. Matthew used more real-time writing on the board, which was usually in response to work through the worksheets and graphic organizers he uses in class. Jennifer, on the other hand, used more question posing. There is very limited research comparing the effectiveness of various types of active learning (e.g., see Connell et al., 2016). This should be reassuring to instructors who are considering using ALS in their classrooms; at this point, the data do not dictate the use of one specific strategy. Instead, instructors can (and should) choose ALS they are comfortable with and/or trained in.

In addition, despite previous research suggesting that instructors without formal training in science education fail to effectively implement ALS (Andrews et al., 2011), our data suggest otherwise. We therefore argue that, rather than focusing on what form of active learning is “best” (until more data of this nature exist), practical consideration should be given to those ALS with which instructors are most familiar and/or comfortable. Using an active-learning strategy that is compatible with one’s teaching style, beliefs, and pedagogical training may be key to how active learning “works.” Andrews and Lemons (2015) have begun discovery research in this area. Making use of qualitative interview data, the authors examined instructors’ use of case study teaching. Ultimately, Andrews and Lemons used their qualitative data to develop hypotheses to be quantitatively tested. Andrews and Lemons hypothesized that instructors’ personal beliefs, as opposed to empirical evidence, drive their use of and persistence in using ALS. The data presented in this paper, along with Andrews and Lemons’ study, clearly demonstrate that many questions related to instructors’ use of ALS and their compatibility with instructors’ teaching styles and beliefs still need to be explored.

In terms of novelty, the psychosocial data from this study suggest that differences in implementation of ALS may impact students’ attitudes, even if the instruction is not designed to explicitly challenge students’ epistemological beliefs. We observed small differences between Matthew’s and Jennifer’s students in their Enjoyment and Real-World Connections. Although the differences were minimal, the present study raises the question of whether different ALS environments (i.e., strategies, frequency of use, and length of implementation), even when not specifically intended to challenge students’ beliefs, can have differential impacts on students’ attitudes.

The cause of this difference may be related to instructional strategy and/or characteristics of the instructor. In terms of instructional strategy, we predict that students may be more familiar with the use of graphic organizers and worksheets (i.e., they used them more in high school) compared with case studies. As a result, the students were more comfortable learning in this way. However, we also acknowledge that instructor characteristics, including personality, age, and gender can influence students’ enjoyment of a class. The literature demonstrates that graphic organizers and worksheets are frequently used in secondary science classrooms as well as in collegiate STEM classrooms, though perhaps less prevalently (e.g., Duran et al., 2009). The research presented herein was conducted in an introductory course composed primarily of freshmen. It is possible, therefore, that students’ enjoyment was linked to their familiarity with graphic organizers and worksheets, and, as such, students in Matthew’s section were able to more easily adapt to the learning techniques used in their class compared with the students in Jennifer’s class.

Prior research demonstrates that instructional strategies that promote metacognition and provide students an opportunity to discuss scientific phenomena are effective in promoting expert-like thinking (Hammer, 1994; Otero and Gray, 2008). It is possible that Matthew’s daily use of graphic organizers and case studies was more effective in assisting students in evaluating their own learning (i.e., metacognitive strategies) compared with Jennifer’s once-a-week use of clicker-based case studies. However, we did not directly test changes in students’ metacognition. Furthermore, differences in attitudes were minimal and only observed in two of seven categories measured.

Research clearly shows that students’ attitudes influence their persistence in STEM (e.g., Adams et al., 2006); therefore, exploring which and how different ALS promote attitudinal and motivational change is an important area of research. We recommend future research exploring how diverse ALS differentially
influence students’ affect and, ultimately, their academic achievement. Specifically, controlled studies need to be conducted to determine what teaching methodologies allow the students to connect their biology course work to their everyday lives. To do this, instructors should be randomly assigned various ALS or the same instructor should implement various ALS in different classes and compare students’ learning gains, attitudes, metacognition, and motivation following instruction. Furthermore, to explore how these changes occur, student interviews and think-aloud procedures should be conducted.

Several limitations are inherent in this study. The sections of the course were taught at different times of the day, were different in size (~100 vs. 250 students), and therefore it is possible that the students who enrolled in these sections were inherently different. To control for possible differences in the sections of this introductory class, we matched students on several demographic and motivational characteristics (Table 2) and, regardless of instruction, only selected students who took an afternoon class. Because of this, we had a smaller sample size between the two instructors than the total number of students enrolled in the two classes. We have also characterized and attempted to account for demographic and pedagogical differences between instructors (Table 1) that might contribute to the possible significant impacts of ALS on student outcomes within the research context. Finally, this is a comparison between two active-learning environments, and it is possible that if there were many more instructors using these strategies, we would see differences in student outcomes.

In conclusion, our data indicate that students learn when ALS, in the form of graphic organizers/worksheets or clicker-based case studies, are present in a classroom. Additionally, it appears that, at the same dosage of active learning, there is no difference in a large-lecture classroom between using graphic organizers/worksheets compared with clicker-based case studies. At this point, instructors do not need to fret over which ALS they incorporate into their classroom but rather need to determine which strategy is feasible for them and allows them to be best equipped to maximize student learning and affect.

ACKNOWLEDGMENTS

The efforts of many individuals and funding sources allowed this project to be completed. First and foremost, we thank “Jennifer” and “Matthew” for allowing us to collect data in their classrooms. This research would have not been possible without their cooperation. We also thank our undergraduate research assistants, Jennifer Le, Rachel Lenn, and Stefani Ronquillo, for their time coding the instructors’ videos. The University of Northern Colorado’s I@UNC Program and Research Dissemination and Faculty Development Program provided funding for this research.

REFERENCES


American Association for the Advancement of Science (2010). Vision and Change in Undergraduate Biology Education: A Call to Action, Washington, DC.


