5-2020

Understanding Learner-Centeredness and Student Engagement in Undergraduate Biology Education

Ashley Barbara Heim

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

UNDERSTANDING LEARNER-CENTEREDNESS AND STUDENT ENGAGEMENT IN UNDERGRADUATE BIOLOGY EDUCATION

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Ashley Barbara Heim

College of Natural and Health Sciences
School of Biological Sciences

May 2020

This Dissertation by: Ashley Barbara Heim

Entitled: *Understanding Learner-Centeredness and Student Engagement in Undergraduate Biology Education*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy in the College of Natural and Health Sciences in the School of Biological Sciences.

Accepted by the Doctoral Committee

____

Emily A. Holt, Ph.D., Research Advisor

____

Lauryn Benedict, Ph.D., Committee Member

____

Teresa Higgins, Ph.D., Committee Member

____

Lori A. Reinsvold, Ph.D., Faculty Representative

Date of Dissertation Defense ____

Accepted by the Graduate School

____

Cindy Wesley
Interim Associate Provost and Dean
The Graduate School and International Admissions

ABSTRACT

Heim, Ashley Barbara. *Understanding Learner-Centeredness and Student Engagement in Undergraduate Biology Education.* Published Doctor of Philosophy Dissertation, University of Northern Colorado, 2020.

The overarching goal of my dissertation research is to better understand how undergraduate students engage in biology. Considering the notable lack of interest in the sciences among undergraduates in recent years, actively engaging more students in biology throughout college could potentially increase their motivation to learn biology and retain more students in science fields. Using both quantitative and qualitative approaches, I sought to discover the dimensionality of learner-centeredness in the biology classroom using a variety of instruments. Outside of the classroom, I aimed to describe college-age adults' learning experiences at informal learning settings such as zoos via development and administration of a novel survey, as well as to discover whether participation in structured or free-choice learning experiences at a zoo related to undergraduates' motivation and interest to learn biology. I generally concluded that learner-centeredness in the college biology classroom is multidimensional, and often, that perceptions of those in the classroom environment as well as the metrics used to quantify learner-centeredness are misaligned. I found that informal learning experiences of biology undergraduates vary widely. Further, we discovered that all students report increases in motivation and interest to learn biology regardless of structure of learning group or academic level—though we cannot say with certainty that a zoo trip was the

cause of these changes. I suggest that both reforming classrooms to be more learner-centered environments and including more learning experiences at informal settings have the potential to more fully engage undergraduate students in biology and improve retention rates of biology majors over time.

ACKNOWLEDGEMENTS

My journey through higher education in the past eleven years has had its highs and lows, but I have always had the support of those around me. To my friends back home in New Hampshire, those here in Colorado, and in all of the places in between: Thank you for always supporting me through my successes and failures over the past decade, bringing humor and laughter to keep my spirits up, and offering words of encouragement and motivation. To my family back in New England: Thank you for supporting my pursuit of a Bachelor's degree, Master's degree, and Ph. D., even if you didn't always understand what that entailed, and for your love of New England sports teams (since I didn't quite feel this during my seven years in Colorado…). To Jen, my twin sister: Thank you for keeping me sane, sending me goofy photos, and allowing me to visit you in Florida multiple times to be a kid again in Disney World and Harry Potter World (and to play with Ravioli). To Megan: Thank you for being by my side in the unpredictable and crazy adventures we know as grad school. To Tanner: Thank you for your love and support over the last year. To Bort the cat: Thanks for being you. To all my prior students: Thank you for giving me the opportunity to engage in my true passion of teaching biology over the past four years; you all made my experience at UNC that much more memorable! To all the faculty and students in the School of Biological Sciences: Thank you all for being amazing, supportive people and for being so encouraging of first-generation students and my biology education research projects. To my committee: Thank you for your advice and feedback during the past several years, and for reading through over 200 pages of my dissertation. And to Emily Holt, my research adviser: Thank you for everything you have done for me over the past four years; I truly could not have asked for a better mentor. You’ve cheered me on in my successes of publishing papers, teaching, presenting at conferences, and being offered a postdoc researcher position at Cornell, among other things. But you’ve also encouraged and motivated me during the frustrating and challenging times of grad school. As your first doctoral student, please know how much your mentorship has meant to me, and that I would not be where I am today without you.

…and Tom Brady: Thank you for all the championships you’ve won since I’ve been in grad school. Please stay in New England.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO UNDERGRADUATE ENGAGEMENT IN BIOLOGY

The overarching goal of my dissertation research is to better understand how undergraduate students engage in biology. Considering the notable lack of interest in the sciences among undergraduates in recent years, actively engaging more students in biology throughout college could potentially increase their motivation to remain in science, technology, engineering, and mathematical (i.e., STEM) disciplines and resolve the so-called “leak” in the STEM pipeline (Barr, Gonzalez, & Wanat, 2008; Chen & Soldner, 2013). Not only could a more learner-centered approach to learning and teaching biology improve retention rates across college biology programs, but may further enhance the authenticity of undergraduates’ learning experiences in the sciences.

Part 1 of my dissertation focuses on gauging learner-centeredness in the biology classroom. The learner-centeredness of a classroom can be characterized by how actively students are engaged in the learning process, and whether the central focus of the classroom is on the instructor or the student (Fahraeus, 2013). There has been a growing emphasis on the implementation of active learning techniques in biology courses—and in STEM fields in general—with a simultaneous shift away from more traditional, passive lectures (Eagan et al., 2014; Ernst & Colthorpe, 2007; Hake, 1998). This is a necessary evolution in how biology courses are taught. Yet, many instructors are resistant to changes in their teaching styles (Allen & Tanner, 2005; Miller & Metz, 2014; Tsang &

Harris, 2016) and frequently students would rather opt for the more convenient uninterrupted lecture in which limited participation and/or critical thinking, if any, is required of them (Covill, 2011; Fox-Cardamone & Rue, 2003; Tsang & Harris, 2016). Despite these initial hesitations, the learner-centered environment serves as a model for enhanced learning and motivation among students, and more student-centered pedagogies have been shown to improve student attitudes and performance in introductory biology courses (Brownell, Kloser, Fukami, & Shavelson, 2012; McCombs, 2000; Miller & Metz, 2014).

Unfortunately, education researchers often disagree on how to most accurately quantify learner-centeredness, and instructors are often unaware of what metrics are most effective for measuring the learner-centeredness of their classrooms. Faculty and student surveys as well as expert observation protocols are frequently used to gauge the learner-centeredness of classrooms. Faculty surveys are often intended to measure affective characteristics of teaching (e.g., McCombs, 2003) or to quantify pedagogical practices and classroom dynamics based on faculty self-reports (e.g., Trigwell & Prosser, 2004). Likewise, student surveys attempt to measure students' self-reported learning experiences, metacognitive strategies, and perceptions of the overall classroom environment (e.g., Biggs, Kember, & Leung, 2001; Entwistle, McCune, & Hounsell, 2002). Trained observers offer a more objective means of quantifying learner-centeredness based on an outside expert's point-of-view. Available observation rubrics measure the quality or quantity of teaching strategies or tasks and student contributions in a classroom (Amrein-Beardsley & Popp, 2012; Ebert-May et al., 2011; Sawada et al., 2002; Smith, Jones, Gilbert, & Wieman, 2013; Wieman & Gilbert, 2014).

While self-reported surveys and observation rubrics are commonly used, no prior studies have compared perceptions of learner-centeredness among students, instructors, and expert observers, nor analyzed whether perceptions among these groups may be misaligned; this calls into question how efficient each instrument may be in capturing learner-centeredness in the undergraduate biology classroom specifically. Ebert-May et al. (2011) concluded that the self-reported teaching practices of nearly 75% of faculty who claimed to implement active learning techniques in their classrooms instead relied on teacher-centered lectures. A more effective means of objectively classifying classroom activities for the common educator could provide a more valid and reliable means of predicting learner-centeredness in undergraduate courses.

The overall aim of Chapter II (Part 1) was to compare student, teacher, and expert perceptions of learner-centeredness in biology classrooms using several valid and reliable surveys and protocols. The overall aim of Chapter III (Part 1) was to measure the learner-centeredness of biology classrooms using DART (Decibel Analysis for Research in Teaching; Owens et al., 2017) and to assess the effectiveness of this instrument for use by everyday practitioners in the classroom. DART quantifies the learner-centeredness of class sessions by estimating the percentage of time dedicated to Single Voice, Multiple Voices, and No Voices (Owens et al., 2017). More specifically, I sought to discover whether a validated metric of learner-centeredness—the Reformed Teaching Observation Protocol (Sawada et al., 2002)—could predict percent Multiple Voice (as estimated by DART), and further, whether external variables (e.g., demographics of students and instructors, classroom characteristics such as room size and enrollment) could also predict percent Multiple Voice.

Part 2 of my dissertation focuses on better understanding how undergraduates learn biology in informal learning settings. Free-choice learning—defined by the autonomy one has in choosing what to learn, for how long to engage in learning activities, and with whom—in informal learning settings may incorporate a variety of learning experiences (NRC, 1996). The National Science Teachers Association broadly describes informal learning environments as those which occur in out-of-school-time settings (NRC, 2009). Further, Hofstein and Rosenfeld (1996) discussed the potentially dichotomous nature of formal versus informal education by noting that many researchers believe these learning experiences must occur in distinct, non-overlapping settings. Many researchers have recently adopted a hybrid definition of informal education, recognizing that free-choice learning experiences can take place in both formal (e.g., schools) and informal (e.g., museums, zoos, etc.) settings. Crane, Nicholson, Chen, and Bitgood (1994) explained that although learning in informal settings can supplement formal learning, free-choice learning is meant to be implemented outside the classroom both in home (e.g., watching television programs or reading books) and in outside the home settings, such as museums, aquaria, and zoos. As the National Research Council (NRC) stated,

> Humans are inherently curious beings, always seeking new knowledge and skills. That quest for knowledge often involves science: from a child's 'Why is the sky blue?' to a teenager's inquiry into the dyes for a new t-shirt; from a new homeowner's concern about radon in the basement to a grandparent's search for educational toys for a grandchild. Each of these situations involves some facet of science learning in [an] informal setting (NRC, 2009, p. 11).

Much informal education research has been conducted within Science, Technology, Engineering, and Mathematics (STEM) fields, due to the scientific nature of most museums, science centers, zoos, and aquariums (MCZAs). The importance of free-choice learning in informal learning settings is elaborated in the National Science Education Standards (NRC, 1996), which highlight the effectiveness of MCZAs in both motivating students to persist in the sciences and increasing their understanding of science outside the formal classroom. Gardner (1991) discussed the influence of informal education within the sciences, suggesting that MCZAs, in general, engage students, increase students' understanding of science, and encourage students to take ownership of their own learning, more effectively than the average science classroom in primary and secondary education. Given the potential benefits of engaging students in informal education opportunities, it is important to consider the learning outcomes and motivations associated with such experiences.

While there is an abundance of research available on free-choice learning in informal learning settings across primary and secondary education, a dearth of knowledge exists regarding the free-choice learning experiences of undergraduates and young adults in informal settings. Informal education research in STEM fields has been almost exclusively conducted at the K-12 level, and while free-choice learning between adolescents and parents as well as programs for youth and the elderly are described within the Venues and Configurations portion of the NRC's *Learning Science in Informal Environments* (1996), the informal learning experiences of college-age adults were not emphasized.

The overall aims of Chapter IV (Part 2) were to use psychometric analyses to analyze the reliability and validity of an instrument that I developed, the Informal Learning Experiences Survey (ILES); to describe young adults' learning experiences at informal learning settings; and to examine which factors predicted the frequency and types of informal learning experiences among members of this age group.

To continue to this exploration of undergraduates' experiences at informal learning settings, I developed a study in which introductory and advanced biology students visited a regional zoo and were randomly assigned to a structured or free-choice learning group. Students in the structured learning group had a specific visitor agenda to follow—enforced by a chaperone—and a structured assessment to complete, while students in the free-choice learning group had autonomy in choosing what exhibits they wanted to visit, for how long, and with whom (given the confines of a college-related field trip). Through questionnaires related to motivation, interest, and self-regulation, the overall aims of Chapter V (Part 2) were to discover whether participation in structured or free-choice learning experiences at the zoo related to undergraduates' motivation and interest to learn biology.

Actively engaging undergraduates in biology courses may provide students more opportunities to think about and discuss biology with their peers (Tanner, 2013). Both improving the learner-centeredness of a class and providing more opportunities for authentic learning in informal settings could stimulate student interest in biology and retain undergraduates in biology degree programs across institutions.

CHAPTER II

COMPARING STUDENT, INSTRUCTOR, AND EXPERT PERCEPTIONS OF LEARNER-CENTEREDNESS IN POST-SECONDARY BIOLOGY CLASSROOMS

This chapter has been previously published in *PLoS ONE*.

Contributions of Authors and Co-Authors

Manuscript in Chapter II

Author: Ashley B. Heim

Contributions: Conceived study topic and design. Organized and analyzed data. Wrote first draft of the manuscript.

Author: Emily A. Holt

Contributions: Conceived study topic and design. Collected initial participant data. Provided feedback on analyses and earlier versions of draft.

Abstract

Learner-centered classrooms encourage critical thinking and communication among students and between students and their instructor, and engage students as active learners rather than passive participants. However, students, faculty, and experts often have distinct definitions of learner-centeredness, and the paucity of research comparing perspectives of these different groups must be resolved. In the current study, our central research question was how do student, faculty, and expert observer perceptions of learner-centeredness within biology classrooms compare to one another? We sampled 1114 students from fifteen sections of a general biology course for non-majors, and complete responses from 490 students were analyzed. Five valid and reliable tools (two faculty; two student; and one expert observer) evaluated the learner-centeredness of each participating section. Perceptions of learner-centered instructors often aligned with those of expert observers, while student perceptions tended not to align with either group. Interestingly, students perceived learner-centered instructors as less learner-centered if they taught at non-traditional times and/or in large-enrollment sections, despite their focus on student learning. Perceptions of learner-centeredness in the biology classroom are complex and may be best captured with more than one instrument. Our findings encourage instructors to be cognizant that the approaches they employ in the classroom may not be interpreted as learner-centered, in the same manner, by students and external observers, particularly when additional course factors such as enrollment and scheduling may encourage negative perceptions of learner-centered practices.

Introduction

Active learning is broadly defined as engaged teaching approaches that encourage critical thinking and communication among students and between students and their instructor (Freeman et al., 2014; McKeachie & Svinicki, 2013; Prince, 2004). Further, active learning contributes to the learner-centeredness of a classroom, which can also be characterized by the level of bilateral learning in a course, and whether students have a role in this process as active learners rather than passive participants (Fahraeus, 2013). While active classrooms tend to share goals of higher cognitive learning and separate the roles of instructors and students in a similar way, they can, on the ground, look very different, depending on the learner-centered practices administered in the classroom.

Experts within education fields have developed these broad descriptions of learner-centeredness and learner-centered practices. However, as Andrews, Leonard, Colgrove, and Kalinowski (2011) noted, the definition of a "learner-centered" classroom is often generated by the instructors or students themselves, generally documented through self-reported survey responses in educational research. It remains unclear to what degree these expert, instructor, and student definitions of learner-centeredness can be interwoven or if they are discrete, potentially diverging perceptions.

Student Challenges with Learner-Centered Classrooms

Learner-centered classrooms reportedly lead to improvements in students' metacognitive abilities, critical thinking skills, and subject knowledge (Armbruster, Patel, Johnson, & Weiss, 2009; Bransford, Brown, & Cocking, 1999; Casagrand & Semsar, 2017; Crouch & Mazur, 2001; Hake, 1998; Holt, Young, Keetch, Larsen, & Mollner, 2015; Shepard, 2000; Knight & Wood, 2005; White & Frederiksen, 1998), and have also been linked with improvements in student performance in the classroom (Armbruster et al., 2009; Freeman et al., 2014; Knight & Wood, 2005; Walker, Cotner, Baepler, & Decker, 2008). Further, increases in student motivation, persistence, self-confidence, and attitudes in science fields have been correlated with learner-centered teaching and learning approaches in STEM (i.e., science, technology, engineering, and technology) courses (Brownell et al., 2012; McCombs, 2000; Miller & Metz, 2014). The multi-faceted, positive impact on students from active learning (Ernst & Colthorpe, 2007; Hake, 1998) is of particular significance in light of the continued leakiness of the STEM pipeline (Chen & Soldner, 2013; Seymour & Hewitt, 1997); perhaps by actively engaging students in STEM courses from the start of their undergraduate careers, instructors can both increase retention rates and ensure a more authentic experience in the sciences for incoming students.

Despite these numerous benefits, many students resist learner-centered pedagogies. University students often have mixed feelings about the use of active learning techniques in lecture (Miller & Metz, 2014; Walker et al., 2008); several studies have reported that students prefer traditional lectures over active learning and consider the former method of teaching more conducive to learning (Covill, 2011; Fox-Cardamone & Rue, 2003; Tsang & Harris, 2016). Herreid and Schiller (2013) noted that students often feel more learner-centered classrooms (i.e. the flipped classroom) require more out-of-class time for reading, homework, etc., than traditional classrooms. Clicker questions or small group discussions in lectures, which require self-directed learning and higher-order thinking of students, have been shown to leave some students feeling frustrated or withdrawn from the course (Felder & Brent, 1996). Similarly, Cooper and Brownell (2016) reported that students of the LGBTQIA community often feel unwelcomed in active learning biology lectures and perceive increased pressure to reveal their identities during the frequent group learning activities characteristic of such sessions. While their study focused on a particular population of students, arguably the transition to a more active classroom likely increases scholastic accountability and social pressure on all students as they are forced into a more collaborative learning environment.

In a study by Watters and Watters (2007), first-year undergraduate biochemistry students reported that they believe effective learning involves information transfer and prefer surface to deep strategies. Therefore, if students understand "learner-centered teaching" as strategies which maximize student learning, which they may erroneously equate with lecture-style presentations, their interpretations of learner-centeredness in the science classroom may be quite skewed from those of instructors and experts. Tsang and Harris (2016), who found that students are unfamiliar with pedagogical practices and the process of learning in general, supports the presence of these student misconceptions. Subsequently, students' negative perceptions of truly learner-centered classrooms and their unwillingness to engage in these practices may be rooted in their misconception that the extra expectations are burdens rather than benefits to them (Weimer, 2002).

Faculty Challenges with Learner-Centered Classrooms

As mentioned above, learner-centered practices may improve student-faculty relations (McCombs, 2000), which consequently improve the overall quality of the classroom environment by providing increased opportunity for discussion amongst the class (Antón, 1999) and shifting the accountability and responsibility of learning from the instructor onto the student (Weimer, 2002). Despite these reported benefits, many instructors remain hesitant to translate learner-centered pedagogies into their current teaching practices, citing lack of support and training (Allen & Tanner, 2005; Miller & Metz, 2014), increased time and effort required to reform a class (Allen & Tanner, 2005; Miller & Metz, 2014; Tsang & Harris, 2016), and loss of "professional identity" (Brownell & Tanner, 2012). Some instructors view the lab component of a course as sufficient engagement and thus fail to incorporate active learning approaches in lecture, demonstrating a form of passive resistance (Brownell et al., 2012; Modell & Michael, 1993). Andrews et al. (2011) argues that the link between active learning and increased student learning gains may be attributed to instructors' pedagogical experience and not the teaching strategy itself. These findings combined with personal ambivalence may deter science faculty from reforming their classrooms, which helps to explain the persistence of didactic lecture (Holt et al., 2015) in the face of contradictory evidence.

However, a gradual shift from traditional lecturing to more active strategies is occurring in undergraduate courses (Eagan et al., 2014), and individual instructors are reforming their classes and experimenting with more learner-centered strategies. Regretfully, approximately 75% of instructors that Ebert-May et al. (2011) surveyed claimed that they used learner-centered practices but in fact used a lecture-based, teacher-driven pedagogy, demonstrating a large disconnect between faculty perceptions and actual teaching practices. This disconnect may derive from the possibility that instructors have their own disparate definition of learner-centeredness compared to students and expert observers, or perhaps because instructors undergo a cognitive shift after pedagogical development that is not necessarily transferred to their actual classroom practices (Guskey, 2002; Huberman, 1981). Dall'Alba and Sandberg (2006) note that, even after educators complete professional development programs, a broad understanding of pedagogical practice is uncommon among participants; the authors further argue that professional development not only incorporates development of skills but knowledge and attitudes as well, which could at least partially explain the aforementioned disconnect between instructors' perceptions of learner-centeredness compared to those of experts. Further, McCombs and Quiat (2002) found that student perceptions tended to be a better measure of learner-centeredness than instructor perceptions and that, additionally, these student perceptions were more aligned with those of trained educational and developmental psychologists rather than the perceptions of course instructors (Daniels, Kalkman, & McCombs, 2001).

Instruments for Measuring Learner-Centeredness

A variety of valid and reliable instruments are available to analyze the learner-centeredness of a classroom (e.g., SETLQ, ATI, RTOP), whether from the perspective of the student, the instructor, or an expert observer. Previous work has used some of these tools to contrast why students learn and how they learn (Biggs, Kember, & Leung, 2001; Ginns & Ellis, 2007; Skogsberg & Clump, 2003; Tiwari et al., 2005), and how the teaching-learning environment influences student approaches to studying and learning (O'Neill & Guerin, 2015; Tudor, Penlington, & McDowell, 2010). Faculty instruments provide teachers formal opportunities for self-reflection and -assessment. Data from these tools may serve as a compass to focus reform efforts to best achieve a student-driven learning environment (Crick, McCombs, Haddon, Broadfoot, & Tew, 2007; Trigwell, 2002; Weinberger & McCombs, 2003). Meanwhile, expert observer protocols are often used to enhance student learning via critiquing and reforming teaching practices from an objective vantage point. Such protocols can quantify the learner-centeredness of instruction in a classroom, providing meaningful feedback to the instructor (MacIsaac & Falconer, 2002; MacIsaac, Sawada, & Falconer, 2001; Sawada et al., 2002).

Many previous studies measure the degree of learner-centeredness of classrooms from just a single perspective: only the student view (Biggs et al., 2001; Ginns & Ellis, 2007; O'Neill & Guerin, 2015; Skogsberg & Clump, 2003; Tiwari et al., 2005; Tudor, Penlington, & McDowell, 2010), only the instructor view (Crick et al., 2007; Trigwell, 2002; Weinberger & McCombs, 2003), or only the expert view (MacIsaac & Falconer, 2002; MacIsaac et al., 2001; Sawada et al., 2002), based on a single instrument; yet, there is a dearth of studies which cross-evaluate student, faculty, and expert perceptions. As students, faculty, and experts often have distinct definitions of learner-centeredness, the paucity of research based on instruments which capture the perspectives of these different groups must be resolved. One exception, Trigwell, Prosser, and Waterhouse (1999), compared faculty and student perceptions with separate faculty (i.e. the Approaches to Teaching Inventory) and student tools (i.e. the Study Process Questionnaire). They found student and faculty perspectives on learner-centeredness generally agreed (Trigwell et al., 1999). In courses where instructors self-reported a more teacher-centered focus on transmitting knowledge, students adopted a more surface approach to learning that subject; in contrast, but less strongly, in courses where instructors self-reported a more student-centered focus on conceptual change, students adopted a deeper approach to learning (Trigwell et al., 1999). These findings were not compared to an expert observer's perceptions of learner-centeredness and therefore may have incorporated bias due to instructors' over-estimation of teaching skills or students' resistance or lack of pedagogical knowledge regarding learner-centeredness.

In another study, Gibbs and Coffey (2004) compared an instructor tool to two student surveys and found that instructors, who were pedagogically trained, tended to believe that they were encouraging deeper learning approaches compared to instructors who received no pedagogical training. While student learning gains improved in courses with pedagogically trained versus untrained instructors, student scores on the "Deep Approach" subscale of a student questionnaire did not significantly increase; in contrast, student learning gains remained unchanged in courses taught by the untrained cohort of instructors (Gibbs & Coffey, 2004). This study suggests that students may be misjudging their learning by performing at a high level but not attributing that success to learner-centered approaches; meanwhile, instructors of their sample who participated in pedagogical training appear more likely to use learner-centered teaching practices and may excel in such aspects of teaching as enthusiasm, organization, and rapport (Gibbs & Coffey, 2004).

The current study is unique in that it used several student and instructor instruments from each perspective within the same classroom, and compared these perspectives to one another in addition to expert perceptions of the same biology classrooms. Redundancy in tools for individual populations can allow us to capture different elements of learner-centeredness, providing a more complete understanding of how learner-centeredness is perceived in the undergraduate biology classroom.

Purpose and Research Questions

In the current study, our central research question was:

Q2.1 How do student, faculty, and expert observer perceptions of learner-centeredness within biology classrooms compare to one another?

Specifically, we wanted to (a) compare subscales within individual student and faculty instruments, (b) compare subscales across student, faculty, and expert observer instruments and describe those relationships, and (c) describe the structure of learner-centered classrooms using multiple instruments. We predicted that different instruments, or subscales within a single instrument, measuring learner-centeredness from a single perspective (i.e., faculty or student) would both linearly and positively correlate. We envisaged that faculty perceptions would generally be disconnected from expert perceptions, as supported by Ebert-May et al. (2011). Contrastingly, we predicted that student perceptions would be more aligned with expert perceptions, as supported by McCombs and Quiat (2002) and Daniels et al. (2001). We also predicted that student perceptions of learner-centeredness would be disconnected from faculty perceptions, supported by Fraser’s (1994) findings that student perceptions of instruction and the overall class environment are more negative than instructor perceptions, even in post-secondary education. We hypothesized that a single-dimension framework, characterized by highly learner-centered at one end and highly teacher-centered at the opposing end, would best describe biology classrooms from various perspectives.

Materials and Methods

Ethics Statement

The procedures for this study were approved by the Institutional Review Boards of Utah Valley University (IRB# 01103) and the University of Northern Colorado (IRB

#932641-1; Appendix A1). Written informed consent was obtained by all participating students and faculty at the beginning of the study.

Participants

We conducted an observational study in introductory biology classrooms at one public post-secondary institution in the western US. While this institution is self-described as “engaged” in its mission, instructors were not considered pedagogical experts. We assumed that the fifteen class sections and nine instructors in our study were representative of average undergraduate biology classrooms, and furthermore, that our results would be applicable to biology courses at other post-secondary institutions.

We sampled 1114 students from fifteen sections of a general biology course for non-majors, and complete responses from 490 students were analyzed (i.e., students who completed both the student surveys administered in this study). While volunteer participation can result in non-response bias, our response rate of 44% is proximal to the accepted average noted in psychological studies (Baruch, 1999) when considering the removal of three course sections from the original data set (n = 244 students enrolled; further described below). Our twelve participating class sections varied by student enrollment (min = 16 students per section, max = 391, mean = 91.4) and class meeting time (1 section was a weekend course, 3 were night classes, and 8 met during the weekday).

Nine instructors taught these fifteen sections during Fall 2013 and Spring 2014; six of these instructors taught two sections during the same semester. One of the participating instructors failed to complete both faculty surveys, and consequentially both of this instructor’s sections were removed from our data set (n = 94 students enrolled).

Additionally, one of the participating instructors voiced concern after completing the faculty surveys regarding their inconsistent interpretation of survey questions; to prevent a lack of validity and reliability in our analyses, we also removed this instructor's section from our data set (n = 150 students enrolled). Our final analyses included twelve sections. The remaining seven instructors had various levels of teaching experience: one instructor had taught for 2-3 years; one for 3-5 years; two for 11-20 years; and three for 21 or more years. Additionally, the population of instructors used in this study included tenured and tenure-track professors, as well as adjunct instructors. Course section numbers used in this paper (1-12) reflect their ranked RTOP score (i.e., section one had the highest RTOP score, while section twelve had the lowest RTOP score), and to protect participant anonymity do not link to actual institutional numbering schemes.

Conceptual Framework

We used five valid and reliable tools (2 for faculty, 2 for students, and 1 for expert observers) to evaluate the learner-centeredness of each section participating in this study. The conceptual framework, or null hypothesis, for our work is a one-dimensional gradient, where a tool or subscale within an instrument falls at either end of a learner- to teacher-centered gradient, concomitantly opposing the other end (Figure 2.1). We expect the student-centered end of our gradient to include classrooms where faculty hold more learner-centered beliefs and focus more on conceptual change in their students, and where students incorporate deeper learning approaches and dedicate more class time to building models and sharing ideas with one another. In contrast, at the opposing end of our gradient, we expect a more teacher-centered classroom to include more non-learner-centered beliefs and be more focused on information transfer by faculty to students, and for students to incorporate more surface learning approaches and rarely interact with the instructor or their peers during class.

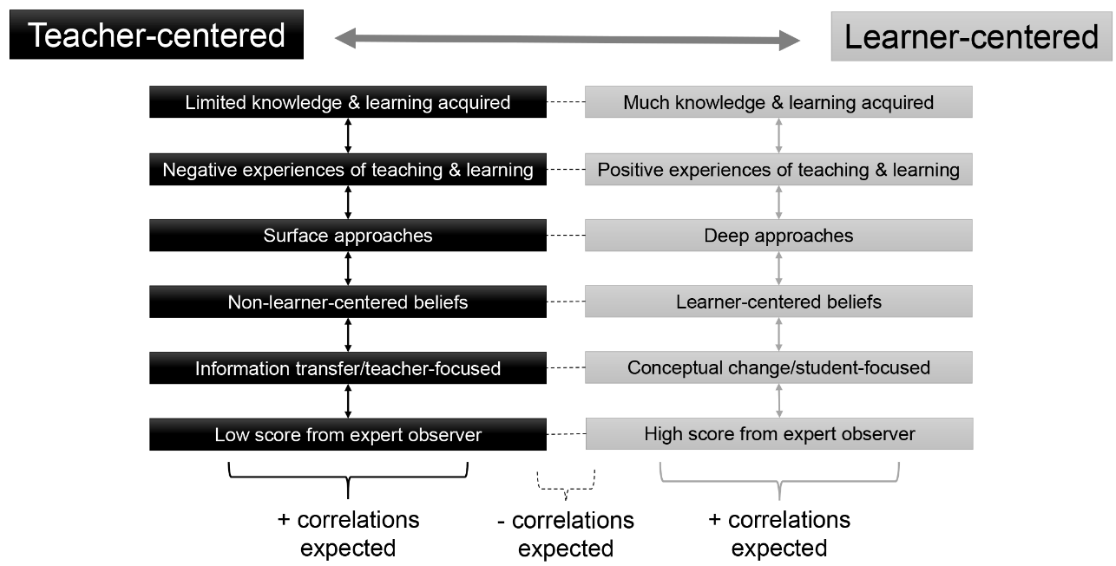


Figure 2.1. The proposed one-dimensional learner- to teacher-centered framework. Examples of student behaviors and instructor practices at the learner-centered end (in gray) juxtapose those that are more teacher-centered (black) at the other end of the framework. Learner-centered descriptors (gray) were expected to positively correlate with each other, while teacher-centered descriptors (black) were expected to positively correlate with each other. Negative correlations (dashed line) were expected between two related but contrasting descriptors, as both would fall on opposite ends of the learner- to teacher-centered framework. For example, deep approaches are more learner-centered, while surface approaches are more teacher-centered; a student that engaged in deeper learning approaches would not be expected to engage in as many surface approaches, or vice versa.

We assumed that subscales or factors of different instruments would overlay onto our conceptual framework (Figure 2.1), and likewise relate to other tools positioned within this framework. If factors, from different instruments or within the same instrument, both attempted to capture learner-centered behaviors, we expected that those factors would positively covary, and fall at the same end of our gradient. Alternatively, we predicted that if one subscale measures teacher-centered beliefs and another measures learner-centered beliefs, they will negatively covary, representing opposite ends of our 1-D framework.

Instruments for Comparing Perceptions of Learner-Centeredness

Nine factors were derived from five published instruments (Table 2.1) to describe learner-centered perceptions in the classroom within our conceptual framework (Figure 2.1). The Assessment of Learner-Centered Practices (ALCP; McCombs & Miller, 2007), a faculty instrument, assessed characteristics of effective teaching, assessment of classroom practices most relative to motivation and achievement, and beliefs and assumptions about learners, learning, and teaching. Two of the three scales within the ALCP measured learner-centered beliefs (LC Bel) and non-learner-centered beliefs (NLC Bel) of faculty. We expected learner-centered beliefs to fall closer to the learner-centered end of the gradient, while non-learner-centered beliefs may fall toward the teacher-centered end of the gradient (Figure 2.1). The Approaches to Teaching Inventory (ATI; Trigwell & Prosser, 2004), founded on research perspectives applied by Marton, Hounsell, and Entwistle (1997), functioned to capture faculty approaches to teaching and learning; the ATI measured information-transfer/teacher-focused (ITTF) and conceptual change/student-focused (CCSF) practices. ITTF practices were expected to overlap with non-learner-centered beliefs at the teacher-centered end of the gradient, while CCSF practices were expected to overlap with learner-centered beliefs near the learner-centered end of the gradient (Figure 2.1).

Table 2.1. **Five instruments for comparing perceptions of learner-centeredness.** Within each student and instructor instrument exists primary and secondary subscales that we used in our study; we indicate the possible score ranges for each subscales and at which end of the learner-centered (LC) gradient a high score on that subscale would capture.

Focus Group	Tool	Primary Subscales	Secondary Subscales	Score range	High score captures which end of the LC gradient?	Citation
Instructor	ALCP	Non-learner-centered beliefs (NLC Bel)	NLC-Bel	5-20	Teacher-centered	McCombs & Miller (2007)
		Learner-centered beliefs (LC Bel)	LC-Bel	5-20	Learner-centered	
Instructor	ATI	Info transfer/teacher-focused (ITTF)	information transfer, teacher-focused	8-40	Teacher-centered	Trigwell & Prosser (2004)
		Conceptual change/student-focused (CCSF)	conceptual change, student-focused	8-40	Learner-centered	
Student	R-SPQ-2F	Deep approaches (Deep)	deep motive, deep strategy	10-50	Learner-centered	Biggs et al. (2001)
		Surface approaches (Surface)	surface motive, surface strategy	10-50	Teacher-centered	
Student	SETLQ	Knowledge & Learning Acquired (KLA)	Knowledge & subject-specific skills (k-skills), generic skills (g-skills), information skills (i-skills)	8-40	Learner-centered	Entwistle et al. (2002)
		Experiences in Teaching & Learning (ETL)	aims, choice, understanding, feedback, assessment, staff, students, interest	25-125	Learner-centered	
Expert	RTOP	N/A	N/A	0-100	Learner-centered	Sawada et al. (2002)

Note. The ALCP only contained primary subscales (NLC Bel and LC Bel), though these factors also served as a proxy for secondary subscale comparisons during our analyses across instruments. Additionally, the RTOP resulted in one average score per class session and we did not further break it down into primary or secondary subscales.

Two student surveys were used to evaluate student learning approaches on a deep or surface level and to better understand the general learning-teaching environment, respectively. The Revised 2-Factor Study Process Questionnaire (R-SPQ-2F; Biggs et al., 2001), based on the original Study Process Questionnaire (SPQ) developed by John Biggs in the 1980s, measured deep and surface approaches. While deeper approaches are motivated by a student's intrinsic interests and desire to maximize meaning, surface approaches are motivated by a student's fear of failure and rote learning strategies (Biggs et al., 2001). We expected deeper approaches to correspond with the learner-centered end of the gradient, while more surface approaches may fall on the teacher-centered end of the gradient (Figure 2.1). The Shortened Experiences of Teaching and Learning Questionnaire (SETLQ; Entwistle et al., 2002) was produced as part of the Enhancing Teaching-Learning Environments in Undergraduate Courses Project and was intended to enhance student achievement via the strengthening of student-instructor relations and of the learning-teaching environment in general (Entwistle et al., 2002). The SETLQ measured six scales, and we focused on two of those scales: student self-reported experiences of teaching and learning (ETL) and knowledge and learning acquired (KLA).

We anticipated that students who self-reported increased learning gains in the classroom (KLA), in addition to having positive teaching and learning experiences (ETL), would cluster near the learner-centered end of the gradient; it should be noted that this is the only pair of subscales from a single instrument that were expected to associate with the same end (i.e. the learner-centered end) of the learner- and teacher-centered spectrum.

The Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2002) quantified the learner-centeredness of instruction within each classroom, as determined by an external observer. The RTOP, originally designed by the Evaluation Facilitation

Group of the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT), allowed trained experts to objectively classify teaching in a classroom on the same learner- to teacher-centered spectrum described above (Figure 2.1). More learner-centered classrooms should earn higher RTOP scores, while more teacher-centered classrooms should earn lower RTOP scores. Sawada et al. (2002) used RTOP to quantify the learner-centeredness of undergraduate science classrooms after instructors participated in professional development workshops.

In the current study, we chose to use RTOP rather than other expert observer tools such as the Classroom Observation Protocol for Undergraduate STEM (COPUS). RTOP requires more rigorous multi-day training to achieve sufficient interrater reliability (Sawada et al., 2002), and contains protocol items that are more aligned with quantification of learner-centeredness in the classroom. Considering expert observer tools, RTOP was the best fit for our research objectives centered on learner-centeredness in the undergraduate biology classroom; per Sawada et al. (2002), RTOP is "standards based, inquiry oriented, and student centered" (p. 1).

Administration and Analysis of Faculty Instruments

Faculty surveys were administered online during the last week of the semester (via www.surveymonkey.com); however, instructors were given up to two weeks to complete the two faculty surveys to maximize response rates. In this study, ALCP (McCombs & Miller, 2007) items were ranked on a 4-level Likert scale and ultimately, answers were categorized into either "learner-centered beliefs" or "non-learner-centered beliefs" (Scales 1 and 3, respectively); scores were then summed based on the system described by McCombs and Miller (2007). The ALCP Scale 2, or "Non Learner-Centered

Beliefs about Learners," was not used in this study, because it focused on personal reflection and emotional aspects of teaching (McCombs, 2003; McCombs & Miller, 2007). We felt that personal beliefs about student performance or persistence may or may not translate into an instructor's pedagogical practices, thus did not cleanly overlay with one end of our framework, as we have defined it. The learner-centered beliefs and non-learner-centered beliefs subscales of the ALCP were not further broken down into secondary subscales as the other instructor and student instruments were.

The ATI consisted of sixteen five-point Likert scale items. Answers were ultimately characterized into one of two pedagogical categories of eight items each based on reported teaching practices: teacher-focused and information transfer-based or student-focused and conceptual change-based (Trigwell & Prosser, 2004). We then summed scores for items in each category. Within the ATI, ITTF can be further broken down into information transfer and teacher-focused and CCSF can be further broken down into conceptual change and student-focused. Hence, an instructor with a high ITTF score would tend to lecture at students more, while an instructor with a high CCSF score would generally focus more on students' understanding of concepts rather than simply transferring knowledge.

Administration and Analysis of Student Instruments

The R-SPQ-2F asked students to respond to twenty items related to attitudes towards and usual methods of studying; the scale for each item ranged from 1 (never or only rarely) to 5 (always or almost always). Main scale scores were categorized into one of two categories and summed: deep or surface approaches (Biggs et al., 2001). Within the R-SPQ-2F, the deep subscale can be further broken down into deep motive and deep strategy, while the surface subscale can be similarly broken down into surface motive and surface strategy. In this case, motive refers to a student's justification for learning and succeeding in the classroom, while strategy refers to a student's plan for learning the material in a particular course and how effective they are in doing so.

Although the SETLQ is composed of six sections, we used only two subscales (the ETL and KLA, described above) in this study due to our perception of their direct relevance to learner-centeredness. The ETL asked students to indicate their level of agreement on 25 items, of a 5-level Likert scale, based on their general approaches to studying and learning. The KLA asked students to respond to eight items regarding their perceptions of what they had learned in the course (i.e., Introductory Biology); the scale for each item ranged from 1 (very little) to 5 (a lot). Scores for each subscale were calculated by summing item responses in a given subscale. Within the SETLQ, the ETL can be further broken down into Aims and congruence (aims), Choice allowed (choice), Teaching for understanding (understanding), Set work and feedback (feedback), Assessing understanding (assessment), Staff enthusiasm and support (staff), Student support (students), and Interest and enjoyment (interest), while the KLA can be further broken down into knowledge and subject-specific skills (k-skills), generic skills (g-skills), and information skills (i-skills).

Both student surveys were administered online during the last week of the semester (via www.surveymonkey.com) and students were given a week and compensated 1% of their final grade to complete them. Additionally, at the beginning of the semester, students were administered a demographic questionnaire and a critical thinking survey used for another study (Holt et al., 2015). The demographic survey included seven questions and collected the ethnic and educational backgrounds of the student participants. Demographic information was available for 94% of students in the current study.

Collection and Scoring of Expert Instrument

During Fall of 2013 and Spring of 2014, 65 classroom sessions of the 12 introductory biology sections were recorded. Filming days were generally selected at random, and each section was recorded between four to eight times during semester, usually without advance notice to the instructor. Three to four usable videos from each section were randomly selected to evaluate using the RTOP. We expected that analyzing multiple class sessions would provide a more comprehensive range of pedagogical strategies the instructors employed throughout the semester, hence representing a more genuine measure of learner-centeredness in the classroom. The RTOP is a tool, considered both valid (Sawada et al., 2000; Sawada et al., 2002) and reliable (Amrein-Beardsley & Popp, 2012; Marshall, Smart, Lotter, & Sirbu, 2011), which quantitatively measures the learner-centeredness of instruction in a classroom. In this study, videos were independently rated by at least two trained raters and inter-reliability was high (see Holt et al., 2015).

Three scales exist within the RTOP, including lesson design and implementation, content, and class culture; items within each scale (25 total) were ranked on a scale from zero (absent) to four (present; Sawada et al., 2002). The summed scores from the 25 items results in an RTOP lesson score ranging from 1-100. Two trained raters (Holt et al., 2015) independently scored each class session. Each score was categorized into one of five RTOP levels (Ebert-May et al., 2011; Sawada et al., 2002). If both raters' scores categorized the same class session into the same RTOP level, the scores were averaged; however, if two scores for a single class session fell into different RTOP levels then an additional tie-breaker rater was used and the two scores sharing an RTOP level were used and averaged. Multiple class session RTOP scores for each section were averaged into a single score. We could not use the natural scales within RTOP, since our final RTOP score for each section represented an average among several raters and class sessions.

Data and Analyses

Cronbach's reliability analyses for each scale were calculated in SPSS (IBM Corp., 2013). From the nine subscales representing three perspectives (student, instructor, and expert observer), we created five data matrices which were used in multivariate analyses. We initially created two sets of these five data matrices; one set used section (n = 12) as the sample unit and the other set used individual students (n = 490) as the sample unit. For each set, the first two matrices included student data: student primary subscales (4 factors) and student secondary subscales (15 factors). The next two matrices included faculty data: instructor primary subscales (4 factors) and instructor secondary subscales (6 factors). The final data matrix, RTOP scores (1 factor), represented expert observations of the same classes.

Unfortunately, we found cluster analyses with student as the sample unit were unwieldly in size (i.e., 490 branch tips), not informative, and did not produce identifiable patterns within the cluster dendrograms. Further, the overall patterns in the ordinations and proportion of variance explained was similar using students or sections (i.e., all students within a section averaged) as sample units. We further discovered that secondary subscales in ordination analyses may be more accurate in parsing out perceptions of learner-centeredness with section as sample unit compared to using student responses as sample unit, though we found no difference in comparing primary subscales using section versus student responses as sample units. Particularly in science education, the use of individual student responses as sample units often leads to an inability to distinguish between learning gains due to instructional practices or learning gains due to extrinsic factors (e.g. experiences and backgrounds) of individual students (Theobald & Freeman, 2014). While individual student responses may seem more attractive as a sample unit, they act as pseudoreplicates; therefore, sections as sample units are statistically superior. Results using students as sample units, therefore, are not reported here and all subsequent analyses reflect sections.

Pairwise Pearson correlations of univariate factors were run in SPSS (IBM Corp., 2013). We compared all our factors, including RTOP scores and student and faculty instruments, at either the primary subscale (i.e. ITTF, CCSF, LC-bel, NLC-bel, Deep, Surface, ETL, and KLA; Table 2.2) or secondary subscale (discussed in the Administration and Analysis of Student/Faculty Instruments sections above). Correlations were compared to a null hypothesis of no relationship, and the resulting p-values were compared to a Bonferroni-adjusted alpha of 0.000806 for the primary subscale comparisons (Table 2.2) and 0.000113 for the secondary subscale comparisons. The Bonferroni-adjusted alpha corrected for multiple comparisons to reduce the possibility of measuring false-positive results.

Table 2.2. Pearson correlations between primary instructor subscales, primary student subscales, and RTOP scores across all sections.

		Instructor (ATI)		Instructor (ALCP)		Expert	Student (R-SPQ-2F)		Student (SETLQ)	
		ITTF	CCSF	LC-bel	NLC-bel	RTOP	Deep	Surface	ETL	KLA
Instructor (ATI)	ITTF	1	-0.55	-0.54	0.19	-0.57	-0.16	0.15	-0.17	-0.45
	CCSF		1	0.36	-0.15	0.57	0.11	-0.74	0.81	0.77
Instructor (ALCP)	LC-bel			1	-0.23	0.32	0.18	0.20	-0.16	0.40
	NLC-bel				1	-0.28	-0.02	0.07	0.00	0.26
Expert	RTOP					1	0.60	-0.23	0.26	0.50
Student (R-SPQ-2F)	Deep						1	0.23	-0.18	0.28
	Surface							1	*-0.97	-0.37
Student (SETLQ)	ETL								1	0.53
	KLA									1

Note. (*) indicates a significant relationship at the corrected alpha of 0.000806, compared to a null hypothesis of no relationship.

We ran non-parametric multidimensional scaling (NMS) analyses, using a Euclidean distance measure, in PC-ORD 7 (McCune & Mefford, 2016) to identify multivariate gradients in perceptions of learner-centeredness and visually capture how various perceptions overlap. We chose to use the student primary subscale data as the main matrix upon which to build ordinations and all other data as secondary matrices to investigate after-the-fact relationships with this matrix. We selected the student matrix, instead of the faculty matrix, because it represented a larger sample (i.e., 490 students vs. 7 faculty members); further, students are the natural center point of a learner-centered classroom, so we wanted to align all other perspectives to theirs.

Mantel tests, or multivariate correlations, between all five matrices (i.e., instructor to student, instructor to expert, student to expert; including both primary and secondary subscales) were also conducted in PC-ORD 7 using Euclidean distances. Lastly, cluster analyses using Ward's minimum variance method to estimate the expected number of clusters (based on a Euclidean distance measure) were run in PC-ORD 7 to further analyze how alike course sections were based on instructor versus student perceptions. Cluster analysis is a multivariate technique that separates data into meaningful groups (or clusters) based on overall relatedness; hence, items that cluster together are more related than items that do not cluster into the same group (Kaufman & Rousseeuw, 2009).

Results

Participating Students, Instructors, and Class Sections

Of the 490 students in our sample who fully completed the demographic portions of the student surveys, 30.8% (151 students) were freshmen, 43.3% (212) were sophomores, 19.6% (96) were juniors, 5.1% (25) were seniors, and 1.2% (6) were post-baccalaureate. The mean self-reported grade-point average within this student population was 3.3 on a 0.0-4.0 scale, while the mean ACT score was 22.9. The majority of participants (79%; 389 students) were Caucasian; 9% (46) were Latina/o; and 12% (55) were other ethnicities. Students, on average, had taken 1.2 biology courses in high school and 0.2 biology courses at the college level.

On average, students scored a 28.7 on the Deep subscale of the R-SPQ-2F (min = 10, max = 50; overall scale reliability $\alpha = 0.842$) and a 28.2 on the Surface subscale of the same survey (min = 10, max = 50; overall scale reliability $\alpha = 0.805$). On the SETLQ, students scored an average of 82.3 on the experiences of teaching and learning (ETL) subscale (min = 25, max = 125; overall scale reliability $\alpha = 0.960$) and a 26.4 on the knowledge and learning acquired (KLA) subscale (min = 8, max = 40; overall scale reliability α = 0.899). It should be noted that the minimum and maximum values reported for each subscale describe both actual student scores and the range of each subscale.

Instructors, on average, scored a 23.9 on the information-transfer/teacher-focused (ITTF) subscale of the ATI (min = 17, max = 33; overall scale reliability α = 0.727) and a 27.1 on the conceptual-change/student-focused (CCSF) subscale of the same survey (min = 20, max = 32; overall scale reliability α = 0.534). Low reliability of the CCSF subscale is most certainly skewed by the incredibly low reliability of the SF portion of the subscale (α = 0.090) rather than the CC portion of the subscale (α = 0.634). For both of the ATI subscales, scores can range from 8-40. The average instructor score on the learner-centered beliefs subscale of the ALCP was 15.6 (min = 11, max = 20; overall scale reliability α = 0.781) and on the non-learner-centered beliefs subscale was a 12.6 (min = 9, max = 16; overall scale reliability α = 0.381). For both of the ALCP subscales, scores can range from 5-25. Low overall scale reliability for instructor subscales could be attributed to the low instructor sample size (n = 7). The average RTOP score among instructors was 40.1 (min = 32.17, max = 54.42), for which scores can range from 0-100.

Pairwise Univariate Correlations

Primary subscales. Comparing primary subscales (e.g. ITTF, CCSF, LC-bel, NLC-bel, Deep, Surface, ETL, and KLA) and RTOP across sections via Pearson correlations (Table 2.2), the strongest negative correlation was measured between ETL and Surface ($r = -0.97$; $p < 0.000806$), which represent student subscales from different instruments. We found no strong positive correlations between primary subscales ($p > 0.000806$; Table 2.2) across sections.

Secondary subscales. Secondary subscales identified above in the Methods were also compared across sections via Pearson correlations. We identified no strong negative nor positive correlations between any secondary subscales ($p > 0.000113$) across sections.

Multivariate Trends Among Instruments

Ordinations. In analyzing average student responses of primary subscales (e.g. Deep, Surface, ETL, and KLA) across our twelve sections, the final stress for a two-dimensional solution was 1.2067 ($p = 0.0199$), with a final instability of <0.001 after 52 iterations (Figure 2.2). We rotated this ordination by the strongest variable, ETL (353 degrees), to load it on a single axis. Axis one explained 96.3% of the variance and axis two explained 3.3% of variance in student primary subscale scores. ETL ($r = 0.99$) and KLA ($r = 0.83$) explained most of the positive end of axis one, while the opposing end of axis one was associated with Surface approaches ($r = -0.60$). Axis two opposed Deep approaches ($r = 0.91$) and somewhat KLA scores ($r = 0.57$) at the positive end and Surface approaches ($r = -0.67$) at the negative end. The positive end of Axis 1 was characterized by learner-centered strategies, while the negative end was indicative of non-learner-centered strategies. Similarly, the positive end of Axis 2 was characterized by learner-centered motives, while the negative end was indicative of non-learner-centered motives (Figure 2.2).

When student secondary subscales by section were overlaid onto the student primary student subscales ordination, the positive end of axis one was associated with several of the secondary subscales, including those of the ETL (SETLQ): feedback ($r = 0.97$), understanding ($r = 0.97$), choice ($r = 0.90$), aims ($r = 0.90$), interest ($r = 0.82$), staff (r = 0.59), and student (r = 0.58); those of the KLA (SETLQ): k-skills (r = 0.84), i-skills (r = 0.72), and g-skills (r = 0.67); and one from the R-SPQ-2F: deep strategy (r = 0.53). Assess was the only secondary subscale of the ETL that did not strongly correlate with the positive end of axis one (r = 0.35). It should be noted that Deep approaches in the primary subscales above did not strongly associate with axis one, although strong correlations did arise among the Deep secondary subscales and axis one. The opposing end of axis one was only strongly associated with the R-SPQ-2F's surface strategy (r = -0.72). The positive end of axis two was correlated with deep strategy (R-SPQ-2F; r = 0.91), deep motive (R-SPQ-2F; r = 0.88), and g-skills (r = 0.66), while surface motive (R-SPQ-2F; r = -0.74) was the only secondary subscale strongly related to the negative end of axis two (Figure 2.2).

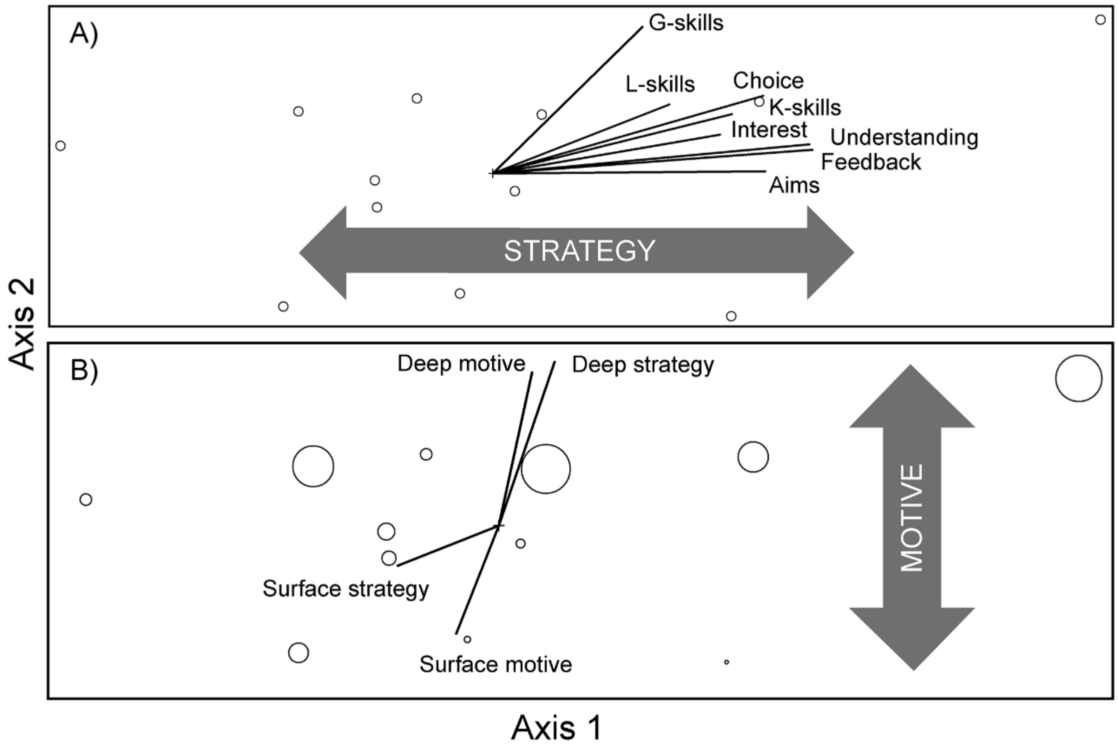


Figure 2.2. Twelve course sections are shown as open circles in student primary subscale space using NMS. (a) Several components of the ETL and KLA positively correlate with Axis 1, the strategy axis. Conceptual change of the ATI also correlated at the positive end of axis one, though was not included in the ordination figure. (b) The Deep and Surface approaches of the R-SPQ-2F associate with the positive and negative ends of Axis 2, the motive axis, respectively. In this panel, the relative symbol size of the 12 course sections are coded by RTOP score; high RTOP scores (i.e., larger circles) correlate with the positive end of Axis 2.

When instructor primary subscales were overlaid onto the ordination of mean student responses per section in primary subscale space, CCSF (ATI) was related to the positive end of axis one (r = 0.63), while no factors were strongly associated (r > -0.5) with the negative end of axis one nor either end of axis two. When instructor secondary subscales were overlaid onto the student primary subscales, conceptual change (ATI) associated with the positive end of axis one (r = 0.61), while no factors were strongly associated (r > ±0.5) with the negative end of axis one nor either end of axis two. The single factor which captured expert perceptions, RTOP, correlated with the positive end of axis two (r = 0.68) but was not strongly associated with axis one. The primary subscales from the second instructor tool, the ALCP, were not strongly associated with either axis ($r < \pm 0.5$) (Figure 2.2).

Multivariate Correlations

Pairwise Mantel tests jointly compared multiple indices of student, instructor, and expert perceptions of the learner-centeredness of participating classes. No significant correlations ($p < 0.05$) existed among class sections based on similarities using primary subscales of instructors and students or RTOP (Table 2.3). Similarly, no significant correlations ($p < 0.05$) existed among class sections based on similarities using secondary subscales of instructors and students or RTOP (Table 2.3).

Table 2.3. Mantel tests between primary and secondary subscale scores. Correlation coefficients and p-values in upper corner compare primary subscale scores, while correlation coefficients in the lower corner compare secondary subscale scores.

	Instructor	**Expert**	**Student**
Instructor	1	p=0.24; r=0.16	p=0.82; r=0.03
Expert	p=0.23; r=-0.16	1	p=0.22, r=0.20
Student	p=0.13; r=0.02	p=0.20; r=0.00	1

Cluster Analyses

To further analyze the relatedness of instructor to student perceptions of learner-centeredness, we compared independent cluster dendrograms based on section-averaged primary subscale responses. Dendrogram nodes were rotated to best align clusters of sections between student and instructor perspectives (Figure 2.3). Some pairs of course sections (i.e., 2 and 4; 11 and 12; 7 and 9; 5 and 6; and 8 and 10) were taught by the same instructor, thus their faculty survey scores are identical. In grouping course sections by student primary subscales (Fig 2.3a), we identified two main clusters with 50%

information remaining. The first student cluster (top cluster; Fig 2.3a) included three course sections (i.e. 2, 12, and 4) in which students tended to have higher ETL, KLA, and deep scores and lower surface scores; this first group was categorized as the more learner-centered group in which learning was based on deep approaches. Interestingly, this cluster also included more of the low enrollment course sections (mean = 57.67 students per section, range = 48-75 students). The second student cluster (bottom cluster; Fig 2.3a) included nine course sections (i.e. 11, 10, 1, 8, 3, 6, 7, 5, and 9) in which students tended to have low ETL, KLA, and deep scores and high surface scores; this second group was categorized as the more non-learner-centered group in which learning was based on surface approaches. Interestingly, this cluster also appeared to include more of the higher enrollment course sections (mean = 102.67 students per section, range = 16-391 students).

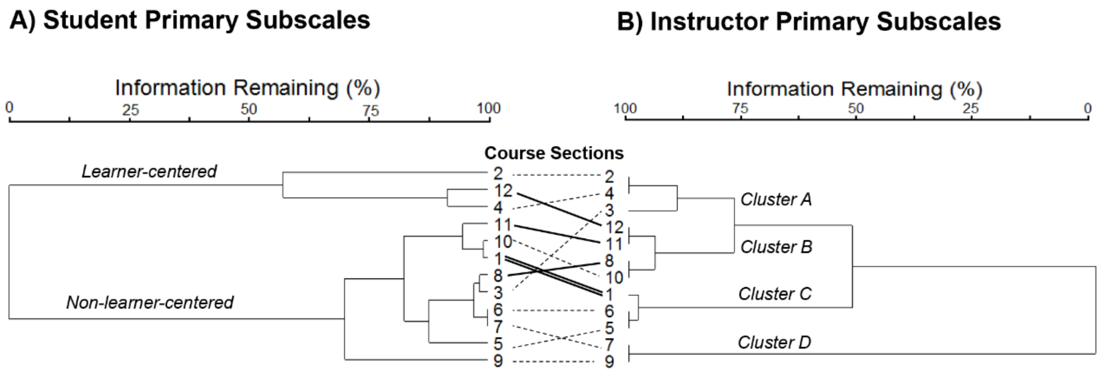


Figure 2.3. Twelve introductory biology course sections independently clustered by student and instructor primary subscales. Sections are clustered by student perceptions in the dendrogram to the left (a), while the same sections are clustered by instructor perceptions in the right dendrogram (b). Identical course sections are connected in the center to aid in visualization of similarities; connector lines patterns denote enrollment size (dashed line ≤70 students, solid line = 71-150 students, bolded double line >150 students [one section, n=391]). In the instructor dendrogram, Cluster A is the true learner-centered cluster; Cluster B is characterized by internal confusion within individual faculty; Cluster C is epitomized by the conflict in perspectives among groups; and Cluster D is the non-learner-centered cluster based on instructor and student perceptions.

In grouping course sections by instructor primary subscales (Fig 2.3b), we identified four main clusters with approximately 85% information remaining. The first faculty cluster (cluster A; Fig 2.3b) included three course sections (i.e. 2, 4, 3) in which instructors were more learner-centered as evidenced by high CCSF scores and three of the top four RTOP scores; interestingly, students also perceived two out of three of these moderately-sized classes to be learner-centered (Fig 2.3a). Cluster A is the only truly learner-centered cluster, where student, faculty, and expert perceptions of learner-centeredness tended to generally align.

The second faculty cluster, cluster B, included four course sections (i.e. 12, 11, 8, and 10) in which instructors were less learner-centered as evidenced by generally higher ITTF and NLC-bel scores; however, sections twelve and eleven had average to high

CCSF and LC-bel scores while sections eight and ten had average CCSF and LC-bel scores (Fig 2.3b). The high CCSF scores in sections twelve and eleven are attributed to high conceptual change scores, as student-focused scores were quite low in these sections. Interestingly, the single instructor of these two sections had more than twenty years of teaching experience and earned relatively low RTOP scores. So while this instructor may have identified with the ideas of learner-centeredness in theory, they may not have put this theory into practice while teaching the sessions we observed. Notably, the instructor of sections 8 and 10 had little teaching experience, which likely influenced their counterintuitive perception of their own teaching as both teacher-focused and student-centered. Students within cluster B perceived these classes to be non-learner-centered, excepting for section 12, in which students perceived the class to be highly learner-centered (Fig 2.3a). Generally, students and experts agreed that the sections in cluster B were non-learner centered, while these instructors expressed mixed views of which end of the spectrum their teaching occupied. Three of the four sections in this second cluster had the greatest student enrollments, excepting section 10, which was closer to the average.

Faculty cluster C included three course sections (i.e. 1, 6, and 5), where instructors had low ITTF scores and high CCSF and LC-bel scores (Fig 2.3b). Cluster C epitomized the conflict in perspectives among groups; while these instructors ranked themselves as highly learner-centered, their students ranked all three of these course sections as non-learner-centered (Fig 2.3a), and experts rated section 1 as learner-centered yet the other two as transitioning to learner-centered. While section 1 had the largest enrollment (n = 391) and was taught during weekday mornings, sections 5 and 6 had the smallest enrollments (n = 16 and n = 30, respectively) and were taught at more non-traditional times (on weekday evenings and weekends, respectively).

Finally, faculty cluster D included two course sections (i.e. 7 and 9) in which the single instructor who taught both sections had high ITTF scores and low CCSF and LC-bel scores (Fig 2.3b); these two courses represented the most teacher-centered faculty cluster. Students agreed that these sections were non-learner-centered, and experts scored them as in the low range of the RTOP level 2, just above teacher-centered.

While most course sections within the instructor and student dendrograms could be roughly aligned (as denoted by straight or nearly straight dashed lines connecting Figs 3a and 3b), some misalignments of sections based on instructor primary subscales versus student primary subscales occurred. Expert scoring of the learner-centeredness of these sections, also did not necessarily agree with these designations. Additionally, student primary subscale scores of two sections taught by the same instructor were never more similar to one another than they were to scores from other instructors' sections. For example, though sections 11 and 12 were taught by the same instructor, students perceived section 11 as non-learner-centered and section 12 as learner-centered.

Discussion

How Did Subscales Within and Among Student Instruments Compare?

Most of the primary and secondary subscales of the SETLQ positively and linearly correlated, suggesting that students' positive experiences with learning coincide with their perceived knowledge gained. Entwistle (2008) reported similar associations linking classroom experiences with conceptual understanding and knowledge acquired, and noted that the extent of conceptual understanding or knowledge acquired may also be influenced by a student's decision to approach learning at a deep or surface level. While students' strategies and motives for learning were orthogonal in our analysis, Deep and Surface approaches fell at each opposing end of both ordination axes (Fig. 2). The ETL, KLA, and deep strategies fell together at the learner-centered end of the same axis, axis one. This alignment supports the idea that students who report having more positive classroom experiences and highly valuing course content tend to adopt deeper strategies (Floyd, Harrington, & Santiago, 2009). The alliance of the two student surveys administered in this study suggests that the R-SPQ-2F and SETLQ can be used in conjunction with one another to capture students' strategies and motives, experiences in teaching and learning, and knowledge acquired on a learner- to non-learner-centered gradient.

How Did Subscales Within and Among Instructor Instruments Compare?

In univariate contrasts, neither primary nor secondary subscales of the ATI significantly related to one another, in agreement with prior studies (Lasry, Charles, Whittaker, Dedic, & Rosenfield, 2013). Surprisingly, the two subscales of the ALCP did not significantly correlate to one another or any of the other faculty scales. Affective aspects of teaching, measured by the ALCP, were likely not captured with the other instruments we used in our study. Low reliability of ALCP scales within our sample population, particularly for the non-learner-centered beliefs subscale, suggests this tool is not reliable with our instructor population thus may be ineffective to measure our desired factor, learner-centeredness. The lack of alignment we observed between the ATI and

ALCP, at least the learner-centered beliefs scale that was moderately reliable, might suggest there is an additional dimension of learner-centeredness among instructors that the ATI did not capture, and which may reflect affective rather than practical aspects of learner-centered pedagogy.

Is Learner-Centeredness Best Represented as a One-Dimensional Gradient?

We found student perceptions of learner-centeredness in introductory biology classrooms are multidimensional (Figure 2.2). Most of the variance among class sections, however, is loaded along one gradient, in line with our original hypothesis that perceptions of learner-centeredness would fall on a single-dimensional framework with two opposing ends. In the student survey, the R-SPQ-2F, the two secondary subscale factors (i.e., strategy and motive) became important but separate factors with surface and deep ends, which defined our two ordination gradients. While strategy represents one’s process or plan for learning, and motive represents one’s orientation for learning, it is important to keep in mind that multiple motive-strategy combinations may be possible; for example, a student may have deep motives but surface strategies for learning a topic (Chiou, Liang, & Tsai, 2012).

We defined Axis 1 as the strategy gradient. Positive experiences of teaching and learning, increased knowledge acquired, deep strategies, and conceptual change describe the learner-centered end of this axis, whereas surface approaches describe the opposing, teacher-centered end (Fig. 2). While the various primary and secondary subscales measured in this study did not covary using linear, univariate analyses, many of the subscales did overlay when viewed in multidimensional space; all subscales on Axis 1 (i.e. KLA, ETL, conceptual change, and deep strategies) aligned as predicted (Fig. 1). The fact that LC-beliefs did not correlate with these other learner-centered measures may suggest that the ALCP is capturing an additional dimension of learner-centeredness (e.g., perhaps one more focused on affective aspects of instruction). Further, though conceptual change and student-focused comprised the CCSF subscale of the ATI, student-focused did not align with other measures of learner-centeredness. Elsewhere, secondary science teachers who intended to teach toward conceptual change rather than based on information transfer often were not able to implement student-focused practices into their lessons (Tabachnick & Zeichner, 1999) which might explain the disconnect we measured between conceptual change and student-focused of the CCSF in the current study. Moreover, we also cannot overlook the considerable unreliability of the SF subscale in our sample, which likely disrupted any potential underlying trend.

We labeled Axis 2 as the motive gradient. At one end of this gradient, students expressed deep motives and strategies for learning and increased general learning skills, and experts perceived these classrooms as highly learner-centered. Surface motives defined the opposing end of this gradient (Fig. 2). Sambell, Brown, and McDowell (1997) noted that even in a learner-centered environment, a student may not adopt deep learning strategies if he or she is not motivated to engage in high-quality learning. However, students in a classroom are reportedly more motivated to succeed if they perceive that they have some control of their learning (Pintrich, 2003). Further, alignment of expert and student perceptions of learner-centeredness has also been reported previously, including the correlation of high RTOP scores with student conceptual gains and classroom collaboration in a learner-centered course (MacIsaac & Falconer, 2002).

In its entirety, Axis 1 (i.e. the strategy gradient) explained substantially more variance in student scores; thus, may be more informative of students' perceptions of learner-centeredness than Axis 2 (i.e. the motive gradient). While many have discussed the close relationship between conceptions of learning and approaches to learning (Biggs et al., 2001; Dart et al., 2000), others have argued that the interplay between conceptions of learning, approaches to learning, and extraneous factors such as culture is more complicated than a simple causal relationship (Lee, Johanson, & Tsai, 2008; Tsai, 2004). While the design of our study cannot infer causation, the strategies students use correlate with a perceived gain in learning (in the form of ETL and KLA scores), but motive is uncoupled from strategy. Though some prior studies have reported that students engaging in deep strategies may not always possess deep motives for learning in a particular course, and vice versa (Chiou et al., 2012), other studies have discussed the strong coupling of deep intrinsic motives and strategies among undergraduate students (Richardson & Newby, 2006). Further, students may perceive their strategies and motives as quite separate entities in the learning process (Chiou et al., 2012), which could be related to the idea that students' conceptions of learning (e.g. motives) may influence their approaches to learning (Edmunds & Richardson, 2009; Marton & Säljö, 2005), whether deep or surface.

Are Two Dimensions of Learner-Centeredness Enough?

Instructor perceptions of learner-centeredness, as measured by the CCSF and CC secondary subscale of the ATI, agreed with student perceptions and fell along the strongest gradient of learner-centeredness, the strategy gradient (Figure 2.2). While instructors in our sample may desire learner-centered outcomes in their classes (i.e., high

CC), some do not engage in the necessary pedagogy to ensure a learner-centered class (i.e., high SF). The paradox of conceptual change in the absence of student-focused learning has been discussed by others in the context of limitations of the original conceptual change model—mainly, that there was too much focus on the instructor's role, rather than the student's role, in facilitating conceptual change in the classroom (Allen & Tanner, 2005; Beeth, 1998; Martin, Mintzes, & Clavijo, 2000; Wandersee, Mintzes, & Novak, 1994). A class based largely on conceptual change is perceived by our sampled students as a class requiring deep strategies and promoting positive learning experiences and increased knowledge and learning. Interestingly, Trigwell et al. (1999) found that student-focused instructors were more likely to encourage deep learning approaches, which our data did not support since high CCSF scores in the current study were mainly driven by the conceptual change secondary subscale rather than the student-focused one. The student-focused subscale was not strongly correlated ($r < 0.50$) to either student gradient, which may suggest additional dimensionality was perceived by instructors but not by students.

Similarly, the two subscales of the ALCP and the ITTF scale of the ATI did not associate with either gradient that students identified as learner-centered. This lack of relationship between the ALCP and other subscales within this study lends more evidence for the multi-dimensional framework of learner-centeredness, even beyond the 2-D model identified in our student ordination (Figure 2.2), rather than the one-dimensional framework described by our null hypothesis. The ALCP, as an example, describes faculty affect that may represent its own separate dimension of learner-centeredness with no relation to the motive and strategy gradients we identified. While prior studies have found strong associations between affective traits of teachers and student outcomes (Roorda, Koomen, Spilt, & Oort, 2011), affective measures of instructors have not historically been linked to instructor and student perceptions of learner-centeredness, as was done in this study by using multiple tools to quantify perceptions of each group.

How Did Subscales Across Student, Faculty, and Expert Observer Instruments Compare?

All univariate and multivariate linear correlations showed no relationships among the student, faculty, and expert instruments, which suggests a disconnect across the subscales of these instruments. However, using data reduction and agglomeration techniques (i.e., ordination and cluster analysis), we were able to identify some overlap in learner-centered perceptions. We found that expert and faculty perceptions mostly align based on cluster analysis; that expert and student perceptions align along the motive axis of the ordination; and that student and faculty perceptions generally do not agree, with the exception of the conceptual change subscale correlating with the learner-centered strategy end of axis one within the ordination.

Similar to our original hypothesis, as guided by work from Ebert-May et al. (2011), our univariate contrasts suggested that expert perceptions of learner-centeredness (i.e. RTOP scores) generally did not relate to faculty perceptions, though our cluster analyses suggested that instructors who perceived their practices and beliefs as learner-centered often taught course sections that were more learner-centered based on expert opinions. Additionally, RTOP scores only associated with the weaker of the two student ordination axes, suggesting that experts' perceptions of the classroom learner-centeredness more closely aligned with students' perceptions of motives rather than strategies. Finally, in agreement with previous work (Fraser, 1994), student and faculty perceptions of learner-centeredness were disconnected in all analyses with one exception (i.e., CC subscale positively associating with the student strategy gradient). Our findings contradict the general agreement between student and instructor perceptions identified by Trigwell et al. (1999) using several of the same instruments administered in the current study, though Trigwell and others noted the small sample size that included only one field of study (i.e., physical science) warranted caution in interpreting the results. Likewise, our study included a relatively small sample (n = 12 class sections) restricted to a single discipline (i.e., biology), which may also contribute to the lack of agreement between our work and Trigwell and others (1999).

Instructors in our study appear to perceive additional dimensions of learner-centeredness that students do not (i.e., measured by the subscales of ALCP), perhaps dimensions based more on affective aspects of teaching and learning. Sutton and Wheatley (2003) discuss the emotional process as relevant to teaching, including how emotional expression and subjective tendencies of teachers may vary during instruction. The ALCP may incorporate this more affective dimension of learner-centeredness, though this dimension could not be adequately detected or aligned with other factors in the current study.

Our finding that RTOP did not associate with the strategy axis of the ordination (i.e., Axis 1) suggests that student strategies do not relate to observable classroom environment and behaviors. As mentioned above, students engaging in deep strategies may not always possess deep motives for learning in a particular course (Chiou et al., 2012). Perhaps the deep motives that many students fostered in the current study were influenced by positive aspects of the classroom environment such as group discussions with peers and a supportive instructor (Rocca, 2010), though these motives may not have necessarily reflected students' strategies to learn biology.

Are Perceptions of Learner-Centeredness Biased by External Factors?

In our sample, we found that the combination of low enrollment courses (i.e., less than or equal to 70 students) with high RTOP scores (i.e., greater than 40) could be viewed as highly learner-centered by both students and faculty. However, in classes where experts and faculty aligned as highly learner-centered yet were either very high enrollment (i.e., greater than 150 students) or taught during non-traditional times (evenings or weekends), students rated these sections as teacher-centered. Differential student success has elsewhere been tied to course scheduling; specifically, students in morning classes outperform students in non-morning classes (Kantartzi, Allen, Lodhi, Grier IV, & Kassem, 2010). Likewise, college science instructors often anecdotally feel that class size is a limitation in implementing more learner-centered or inquiry-based techniques in the lecture (Brown, Abell, Demir, & Schmidt, 2006). Our data empirically suggest that even if a class looks and feels learner-centered, external barriers (i.e., time of day, class size) may limit this perception by students.

Prior studies have concluded that learner-centered practices can be implemented effectively in large enrollment science courses (Allen & Tanner, 2005; Armbruster et al., 2009; Deslauriers, Schelew, & Wieman, 2011). However, our findings demonstrate that while faculty and experts perceive some larger enrollment course sections as learner-centered, students fail to perceive this learner-centeredness when enrolled in these large classes themselves. The tendency of students to perceive larger classes as more teacher-centered in the current study is similar to the trend described by Ebert May et al. (2011) and Murray and MacDonald (1997), though in these prior studies, instructors and experts, rather than students, perceived larger classes as more teacher-centered.

Conclusions

Our sample of introductory biology classrooms clearly implies that learner-centeredness is multidimensional and is more complex than a simple dichotomous learner- versus teacher-centered relationship. The alignment of student, instructor, and expert perceptions of learner-centeredness or teacher-centeredness was generally inconsistent across sections of this non-majors biology course. Broadly, expert opinions tended to agree with instructor and student perceptions independently, while students' perceptions mostly differed from those of faculty. Regretfully, the classroom experience for students can be negatively influenced by external factors, including enrollment size and time of lecture. Future directions of this research should consider interventions to better align perceptions of learner-centeredness in the biology classroom, specifically focused on large or non-traditionally timed courses. Perceptions of learner-centeredness in the biology classroom are complex, and can be more completely measured and interpreted with more than one instrument. Our findings encourage instructors to be cognizant that the approaches they employ in the classroom may not be interpreted as learner-centered, in the same manner, by students and external observers, particularly when additional course factors such as enrollment and scheduling may encourage negative perceptions of learner-centered practices.

CHAPTER III

THE PRACTITIONER’S PANACEA FOR MEASURING LEARNER-CENTEREDNESS?

Contributions of Authors and Co-Authors

Manuscript in Chapter III

Author: Ashley B. Heim

Contributions: Conceived study topic and design. Organized and analyzed data. Wrote first draft of the manuscript.

Author: Emily A. Holt

Contributions: Conceived study topic and design. Collected initial participant data. Provided feedback on analyses and earlier versions of draft.

Abstract

The Decibel Analysis for Research in Teaching (DART; Owens et al., 2017), a sound-based metric of learner-centeredness, is highly accessible, requires no training, and can be conducted with minimal classroom observations; yet, DART has not been evaluated in comparison with other validated metrics or in consideration of potentially confounding classroom characteristics (e.g. enrollment, classroom size, number of doors). We analyzed recordings from 42 class sessions of an undergraduate biology course with DART, the Reformed Teaching Observation Protocol (RTOP), and nine classroom characteristics. We found that enrollment was the best single predictor of the DART output of learner-centeredness, percent Multiple Voice.

Introduction

What is Learner-Centeredness and the Challenges in Measuring It?

Learner-centeredness is characterized by how actively students are engaged in the learning process as they interact with their peers and instructor (Fahraeus, 2013). Often, but not always, active learning is necessary to foster a learner-centered classroom (Cattaneo, 2017). Learner-centeredness has many suggested benefits for students, including lower failure rates (Freeman et al., 2014), improved student performance (Armbruster, Patel, Johnson, & Weiss, 2009; Freeman et al., 2014; Kahl Jr. & Venette, 2010; Walker, Cotner, Baepler, & Decker, 2008), and increased critical thinking skills, metacognitive abilities, and content knowledge (Bransford, Brown, & Cocking, 1999; Crouch & Mazur, 2001; Hake, 1998; Shepard, 2000). Given these benefits, instructors and researchers have sought reliable measures of learner-centeredness for reflection and to guide teaching reform.

Observation rubrics objectively measure the quality or quantity of teaching strategies or tasks and student contributions in a classroom, thus tend to be more accurate than other learner-centered metrics for education research studies (Cohen & Goldhaber, 2016; Shavelson, Webb, & Burstein, 1986). One of the most heavily used observation protocols, the Reformed Teaching Observation Protocol (RTOP; Amrein-Beardsley & Popp, 2012; Sawada et al., 2002), requires time-intensive training, which precludes its accessibility by practitioners. Even observation protocols that require less intensive training (e.g., Classroom Observation Protocol for Undergraduate STEM, Smith et al., 2013; Practical Observation Rubric To Assess Active Learning, Eddy, Converse, &

Wenderoth, 2015; Teaching Perspectives Inventory, Pratt & Collins, 2000) are still time-intensive to conduct, or cannot be conducted with just a few observations (Measurement Instrument for Scientific Teaching-Observable, Durham et al., 2018). Thus, a more automated method of objectively classifying learner-centeredness in undergraduate courses is necessary for accessible and accurate feedback.

Reformed Teaching Observation Protocol

While the RTOP, with its extensive training requirements, is not accessible to all users, it is also considered the standard in observation protocols for discipline-based education research. RTOP has been used across different science fields, including biology (e.g., Ebert-May et al., 2011, 2015; Gormally, Brickman, Hallar, & Armstrong, 2011; Heim & Holt, 2018), physics (e.g., MacIsaac & Falconer, 2002; Falconer, Joshua, Wyckoff, & Sawada, 2001), and chemistry (e.g., Rushton, Lotter, & Singer, 2011). Additionally, RTOP is versatile across education levels—including K-12 (Kilday & Kinzie, 2009; Sawada et al., 2002; Tarr et al., 2008) and college (Amrein-Beardsley & Popp, 2012; MacIsaac & Falconer, 2002; Ebert-May et al., 2011, 2015; Gormally et al., 2011; Heim & Holt, 2018). Researchers have used this instrument to study both longitudinal changes (Ebert-May et al., 2011, 2015) in classroom teaching practices as well as single time points or multiple RTOP scores averaged for individual class sections (Amrein-Beardsley & Popp, 2012; Heim & Holt, 2018; Rushton et al., 2011). RTOP has been used to inform classroom reform (Gormally et al., 2011; Kilday & Kinzie, 2009; MacIsaac & Falconer, 2002) and for professional development (Ebert-May et al., 2011, 2015; Singer, Lotter, Feller, & Gates, 2011). The breadth and adaptability of RTOP make it an ideal instrument for objectively measuring learner-centered teaching practices, and represent the standard against which other instruments have been compared (Heim & Holt, 2018).

Classroom Sound as a Measure of Learner-Centeredness

Studies suggest that types of classroom learning activities can be categorized based on vocal classroom discourse and sound (Kranzfelder et al., 2019; Li & Dorai, 2006; Wang, Pan, Miller, & Cortina, 2014). Kranzfelder et al. (2019) developed the Classroom Discourse Observation Protocol to characterize teacher discourse moves in an undergraduate biology course. Wang et al. (2014) reported that the Language Environment Analysis system, originally designed for infants and pre-schoolers, can distinguish among lecturing, whole class discussion, and group work in an elementary school math class. Li and Dorai (2006) describe two types of vocal discourse: question-and-answer between instructors and students, and group discussions engaging multiple students.

Owens et al. (2017) developed the Decibel Analysis for Research in Teaching (DART), which analyzes audio recordings from a classroom session to estimate the percent of the session dedicated to active versus passive learning strategies, based on an algorithm which outputs the number of voices (i.e., Single, Multiple, or None) extracted from the recording. For a given audio file, DART outputs waveform visualizations and percent ratios of Single Voice, Multiple Voices, and No Voice for each class session, each with a possible range from 0-100%, with the assumption that Multiple Voice and No Voice correlate most with active learning components of learner-centered classroom practices (Owens et al., 2017).

The DART instrument represents an exciting tool to potentially address the need for a universally available, low-cost method for practitioners and researchers alike to categorize the learner-centeredness of undergraduate science classrooms. To date, no study has compared DART estimates to other standard measures of learner-centeredness to describe its validity in reference to other reliable metrics. While DART is accessible and easy to use, it is unclear if the data it provides overlap with elements of learner-centered practices that prior instruments also measure. Hence, we sought to explore whether DART could provide accurate measurements of learner-centeredness comparable to another available metric, thus clarifying the potential of DART to be used by everyday practitioners in the classroom.

External Factors that Contribute to Learner-Centeredness and Classroom Sound

While our first goal was to investigate the alignment of DART with RTOP, we also sought to explore other potential factors that could affect the noise levels of a classroom that may subsequently bias a sound-based metric such as DART. Specifically, we speculated that physical aspects of the classroom itself and the types and background of the people in the classroom may alter both the sound during a class and its learner-centeredness, biasing estimates from DART.

Classroom characteristics. We predicted that numerous physical characteristics of a classroom could affect its learner-centeredness, but these same characteristics also may contribute to noise, unrelated to the quality and frequency of learner-centered activities. For example, some higher education institutions have redesigned their classroom spaces to support active learning (Harvey & Kenyon, 2013) by moving away from a fixed-seat lecture hall (Oblinger, 2006). Despite these redesigns, large classroom sizes, in terms of both enrollment and square footage, still exist and may limit students' motivation to participate in discussions or activities (Abdullah, Bakar, & Mahbob, 2012), minimize support from instructors (Loh Epri, 2016), and increase challenges in classroom management (Ayeni & Olowe, 2016) and hinder large-scale active learning activities. Ironically, although greater enrollment of students in large lecture halls may increase background noise, high enrollment classrooms may lead to decreased engagement (Bradley, 2005; Seep, Glosemeyer, Hulce, Linn, & Aytar, 2000).

Additionally, movable seating and flexible writing surfaces have been found to support more active learning classroom practices (Lombardi & Wall, 2006; Sanders, 2013). For example, flat seating with movable furniture may be more conducive to learner-centered practices when desks are arranged into small groups for discussion (Park & Choi, 2014). The number of doors and windows in a classroom may also influence student engagement. While some suggest that open doors and windows may act as distractors for students and instructors alike by allowing entry of sound from outside the lecture space (Lei, 2010; Veltri, Banning, & Davies, 2006), others emphasize the importance of windows in maintaining a positive and comforting learning environment (Chism, 2006; Montgomery, 2008).

Student and instructor demographics. Beyond the physical characteristics of a classroom, student and instructor demographics may also influence learner-centeredness and classroom noise. Female students are more likely to vocally participate when they have a female instructor (Cornelius-White, 2007; Fassinger, 1996; Pearson & West, 1991); therefore, instructor demographics can influence class engagement. Reciprocally, student gender may influence how students interact with one another and perform (Eddy, Brownell, & Wenderoth, 2014; Eddy, Brownell, Thummaphan, Lan, & Wenderoth, 2015). Male students tend to participate more than their female counterparts and dominate classroom discussions (Howard & Henney, 1998; Pearson & West, 1991), so a class with more male students may be louder than the same-sized class with a lower male:female ratio. Further, because first-generation, low socioeconomic status students, and older non-traditional students tend to experience more social and academic challenges than traditional students (Bowl, 2001; Crosnoe & Muller, 2014; Schuetze & Slowey, 2002; Wilbur & Roscigno, 2016), students in these populations may be less inclined to engage in discussions or collaborative in-class activities (Pike & Kuh, 2005).

Research Goals and Questions

To our knowledge, no research has yet explored the relationships between recorded sound in a classroom using DART, other valid metrics of learner-centeredness (i.e., RTOP), physical characteristics of the classroom, and instructor and student demographics. Many studies have characterized learning activities from audio recordings in a classroom setting, yet these have almost exclusively been conducted at the K-12 level and have generally been implemented only in classes of small size (Donnelly et al., 2016; Donnelly et al., 2017; Wang et al., 2014), excepting the study conducted by Owens et al. (2017). Specifically, there is a need for an accurate, accessible instrument that can be implemented by everyday practitioners in the college classroom. Thus, our research questions were:

Q3.1 Does a validated metric of learner-centeredness—the RTOP—predict percent Multiple Voice from DART?

Q3.2 Do external variables such as classroom characteristics and demographics of instructors and students predict percent Multiple Voice from DART?

Methods

Ethics Statement

The procedures for this study were approved by the Institutional Review Boards of Utah Valley University (IRB# 01103) and University of Northern Colorado (IRB #932641-1; Appendix A1). Written informed consent was obtained by all participating faculty and students at the beginning of the study.

Participants, Classrooms, and Variables

We conducted this observational study within a non-majors introductory biology course at a public 4-year university in the western United States. Nine instructors collectively taught thirteen sections of this introductory biology course during Fall 2013 and Spring 2014. Our instructor sample included four females and five males.

The thirteen class sections in our study varied by several factors. We coded instructor gender into two categories (Table 3.1). De-identified student demographic information was retroactively obtained from the institution's office of institutional research, including gender, first-generation status, age, and Pell Grant eligibility (used as a proxy for students' socioeconomic status), in accordance with our IRB approval. Unfortunately due to considerable missing data, first-generation status and Pell Grant eligibility were not used in our final models. In our analyses, student gender was represented as the proportion of males in a course section, and student age was represented by the mean age of students in a course section (Table 3.2).

Table 3.1. Categorical predictors of % Multiple Voice in the classroom.

	Predictor	Counts for Category	How does it contribute to learner-centeredness?	How does it contribute to classroom sound?
Demographic	**Instructor gender**	**Males** n=5 **Females** n=4	Female instructors may encourage increased participation among female students. Student gender may also influence teaching and learning practices in a class.	Female students may be more likely to vocally participate when they have a female instructor. In the absence of a female instructor, only a proportion of the class (i.e., males, who generally have deeper, louder voices) may be speaking rather than all students, contributing to an overall noisier classroom.
Classroom Characteristics	**Chair type**	**Fixed** n=6 **Non-fixed** n=5 **Mixed** n=2	Physically larger classroom spaces tend to be louder, making it difficult for students to engage in learner-centered practices. However, some classroom attributes such as movable furniture may be more conducive to active learning practices (e.g. group discussions).	Chairs and tables that are non-fixed may be noisier than their fixed counterparts, as students reposition during class. Physically larger classroom spaces with wooden or plastic furniture and stadium seating tend to amplify noise.
	Table type	**Moveable** n=6 **Fixed** n=7		
	Chair material	**Fabric** n=8 **Plastic** n=5		
	Table connectivity	**Individual** n=7 **Shared** n=6		
	Seat arrangement	**Stadium** n=9 **Flat** n=4		

Table 3.2. Continuous predictors of % Multiple Voice in the classroom.

	Predictor	Minimum	Maximum	Mean	How does it contribute to learner-centeredness?	How does it contribute to classroom sound?
Demographic	**Student age**	16	63	22.7	Older, non-traditional students) may disengage from in-class learning activities more so than other students.	Non-traditional students and females may be less inclined to engage in discussions or collaborative in-class activities and quieter in the classroom overall. Hence, we predicted a greater percentage of non-traditional and female students may contribute to a less noisy classroom due to fewer voices being expressed.
	% female students in a section	32.7	64.2	48.0		
Classroom Characteristics	**Enrollment**	30	391	94.6	Large classroom sizes may make learning more difficult and active learning practices less effective due to physical constraints of the classroom and a high quantity of students.	High enrollment of students in large lecture halls may increase background noise, contributing to a louder classroom.
	Room size (sq ft)	691.0	5173.0	1966.1		
	Number of doors	1	16	4.5	Increased lighting may positively affect students and increase their willingness to engage in active learning exercises, though many doors and windows in a classroom could also lead to higher potential for distractions.	More doors or windows in a classroom may increase classroom noise if used frequently.
	Number of windows	0	16	4.6		
RTOP	**Mean RTOP scores per section**	30.2	54.4	38.8	Higher RTOP scores indicate greater learner centered practices by students and the instructor.	Higher RTOP scores could indicate both a noisier classroom (e.g., lots of interactive active learning occurring) or quieter classroom (e.g., silent reflective/thinking exercises).
	Mean Classroom Culture scores per section (RTOP subcategory)	9.0	26.5	15.8		

Our 13 participating sections were scheduled in 9 locations across the same campus. Classroom characteristics were described by an outside observer or from institutional facilities statistics. The classroom characteristics we captured included square footage, number of doors, number of windows, chair type (i.e., fixed, non-fixed, or a combination of these two types), chair material (i.e., plastic or fabric), table type (i.e., fixed or moveable), table connectivity (i.e., shared with peers or individual), seat arrangement (i.e., stadium or flat seating), and section enrollment (Tables 3.1 and 3.2).

Video Recordings

During Fall 2013 and Spring 2014, 42 class sessions were randomly recorded throughout the semester across 13 course sections. A video recording device was situated on a tripod at the back of the lecture space, and each instructor was instructed to secure a wireless lapel microphone and battery pack to their person. The number of class sessions filmed within each course section ranged from three to four. Generally, the instructor was not given advance notice that their lecture would be video-recorded on filming days. These video recordings were used to analyze: (1) audio recordings with the Decibel Analysis for Research in Teaching (DART) instrument; and (2) video recordings for the Reformed Teaching Observation Protocol (RTOP).

Decibel Analysis for Research in Teaching

We converted all video files to .wav audio files compatible with DART using Audacity (Audacity Team, 2017). We also used Audacity to trim each audio file to limit background noise from before class, after class, or breaks, to ensure that the visualizations and predictions generated by DART were solely based on instructional time. Trimmed audio files were individually uploaded onto the publicly available DART software page (Version 1; sepaldart.herokuapp.com; Science Education Partnership & Assessment Laboratory, San Francisco State University).

In this study, our response variable was percent Multiple Voice predicted by DART for each audio file. For a given audio file, DART outputs waveform visualizations and percent ratios of Single Voice, Multiple Voices, and No Voice for each class session, with the assumption that Multiple Voice and No Voice correlate most with active learning components of learner-centered classroom practices (Owens et al., 2017). The No Voice DART category was not detected in any of our audio recordings, thus our use of Multiple Voice percent alone as a response for learner-centeredness is appropriate.

To ensure the validity of DART, we used human annotation on 17% of the data to measure the accuracy of DART, according to Owens et al. (2017). We annotated the two class session recordings with the highest percent Multiple Voice, the two recordings with the lowest percent Multiple Voice, and three random recordings with varying 'moderate' percent Multiple Voice output from within our sample. These annotations consisted of two trained annotators independently coding the length of time spent lecturing with question-and-answer, silent working, discussing in pairs or small groups, or other activities not represented as a prior code, using codings for human annotation described by Owens et al. (2017). Our inter-rater reliability, the Pearson correlation between the two raters across the seven video recordings, of 0.96 was high; Cohen's alpha was inappropriate because our data were continuous rather than categorical.

Reformed Teaching Observation Protocol

The RTOP, considered both valid and reliable (Amrein-Beardsley & Popp, 2012; Marshall et al., 2011; Piburn & Sawada, 2000; Sawada et al., 2002), allows experts to objectively quantify learner-centeredness in classrooms based on observations. In our study, we had eight trained raters who differed from the DART annotators, and differing combinations of two of these raters individually scored each of the 46 video-recorded class sessions (Generalizability Coefficient = 0.787; see Holt, Young, Keetch, Larsen, &

Mollner, 2015) and their scores were averaged for each class session. The RTOP is composed of three scales—lesson design and implementation, content, and classroom culture—from 25 items. Items are scored on a scale from zero (absent) to four (present; Sawada et al., 2002), and scores across all items are then summed to calculate a final RTOP score ranging from 0-100. Thus, a higher RTOP score indicates a more learner-centered classroom. In addition to total RTOP score, we also chose to include the score (ranging from 0-20) from the "Classroom Culture: Student/Teacher Relationships" scale in our models, which is a 5-item scale within RTOP focused on student and instructor interactions that we felt might be more relevant for predicting DART due to its potential alignment with learner-centeredness in the classroom.

As multiple video recordings for one course section (i.e., different meetings from the same class) included redundant data for the instructor, students enrolled in the course section, and classroom characteristics, we were cognizant about the inherent pseudoreplication problem within our dataset and sought to minimize its impact. Thus, we ran each of the models described below with a random subset of 13 individual sessions from the 13 class sections; the variance explained by these models changed drastically when an additional predictor variable was included in the model, suggesting that a single-class subset was a poor approach due to the small sample size. All analyses and results below, therefore, represent the full 42 class sections.

Statistical Analyses

Initial analyses included descriptive statistics to describe participants and classroom characteristics, interpret distributions of the data, and assess suitability of potential variables to be included in our models. Bivariate correlations (Pearson correlations for relationships between continuous data) were conducted in SPSS (IBM Corp., 2017) to measure relationships only between significant predictors in our models. We visually inspected scatterplots for the Pearson correlations to ensure that these data were generally linear in nature. The sample units for our data analyses were individual recordings (i.e., n = 42 class sessions) rather than course sections. In recognition of pseudoreplication mentioned previously within our data, we included both instructor and section number in our models to better understand how this redundancy affected our findings.

We used nonparametric multiplicative regression (NPMR) modeling to identify potential predictors of percent Multiple Voice in the classroom. NPMR is a flexible method of regression that allows for complex interactions that are not possible to analyze with general linear regression models (Berryman & McCune, 2006). NPMR models predict quantitative response variables using a smoothing function and Gaussian local mean estimators and are assessed with a leave-one-out cross-validated R^2 (xR^2). Further, predictors in NPMR models are considered multiplicatively; thus, multicollinearity is not a concern when running these analyses. Scree plots incorporating xR^2 and predictor variables of interest were used to select a final model. We ran our NPMR models in HyperNiche (MjM Software, 2009) with medium overfitting controls, deleting all but the best predictors in the final models.

We developed NPMR models to predict the average percent Multiple Voice based on 16 possible predictors. Our full predictor set included 3 demographic predictors (i.e., instructor gender, student gender, student age), 9 classroom characteristic predictors (i.e., chair type and material, square footage of classroom, number of doors and windows in class, section enrollment, table type, table connectivity, and seat arrangement), and both total RTOP score and Classroom Culture scale score from RTOP. We also included instructor identity number and course section as predictors to detect the effect of pseudoreplication.

Results

Human Annotations of Classroom Activities to Test the Validity of the Decibel Analysis for Research in Teaching Tool

Comparing DART output and human annotations of classroom activities (Table 3.3), the majority of time in each classroom session was spent lecturing (with the exception of Classroom Session 7), yet this value does not specifically align with the percent Single Voice output by DART (other than for Classroom Session 1). Even in Classroom Session 7, where nearly 70% of class time was spent in pair or small group discussions as noted by human annotation, the 30.8% Multiple Voice DART output—though the highest value across all recorded sessions—was rather low. However, this inconsistency at the higher end may have been partially due to microphone issues. Various instances of pair/small group discussions observed through annotation of this class were categorized as a single voice by DART, likely because a single voice of the instructor or a student immediately adjacent to the instructor was louder than the overall student oral discourse in the background.

Table 3.3. Comparison of DART output (in the form of % Multiple Voice) and human annotations of classroom activities. Annotations were conducted on the two class session recordings with the highest percent Multiple Voice, the two recordings with the lowest percent Multiple Voice, and three random recordings with varying 'moderate' percent Multiple Voice output from within our sample. The Pearson correlation for each class session represents the agreeability between the two raters' annotations across each of the five annotation categories listed in the table. The difference in DART and Human Annotation highlights where the two measures agree or not and degree of agreement.

		DART Scoring	Human Annotation Scoring (% of time in class session spent performing each activity)					
			Coded as SV[2]		Coded as NV[2]	Coded as MV[2]	Coding Unkn[2]	Pearson r (inter-rater reliability)
Class Session	Difference in DART & Human Annotation[1]	% Multiple Voice	Lecture without Q&A	Lecture with Q&A	Silent working	Pair/small group discussion	Other	
1	0	0	93.86	5.88	0.00	0.00	0.26	0.9986
2	0	0	52.38	31.92	0.00	0.00	15.69	0.7877
3	10.5	10.5	68.72	31.23	0.00	0.00	0.05	0.9911
4	12.2	12.2	85.75	14.25	0.00	0.00	0.00	0.9819
5	20.3	20.3	88.03	9.17	0.00	0.00	2.80	0.9994
6	22.7	22.7	66.83	30.09	0.00	0.00	3.08	0.9786
7	-37.4	30.8	12.94	5.15	0.00	68.20	13.71	0.9967

Additionally, other sessions were inversely mismatched; when DART detected moderate levels of percent Multiple Voice, annotations consisted primarily of lecture and lecture with question-and-answer (Table 3.3, Classes 3-6). Misalignment of DART output with our human annotations suggests that many instances of lecture with question-and-answer included background student discussions beyond the individual student or instructor asking or answering questions. Hence, this may have been a weakness in our activity categories for annotation (i.e., some question-and-answer time may be more active than we expected), or may suggest that DART was able to better parse out background noise and side discussions among students during lecture with or without question-and-answer.

Descriptive Analyses

Percent Multiple Voice across our 42 sampled class meetings, as predicted by DART, ranged from 0% to 30.8%, with a mean of 7.14% across all recordings. Percent Single Voice across recordings ranged from 81.86% to 100%, with a mean of 93.75%. DART did not detect any instances of 'No Voice' in our sample. Across the 42 class recordings, the mean total RTOP score was 38.8 (i.e., teacher-centered lecture with limited demonstrations and student participation). Ranges and means of continuous classroom characteristics and student and instructor demographics are reported in Table 3.2.

Nonparametric Multiplicative Regression

In our NPMR models, the best predictors of percent Multiple Voice based on DART output were enrollment (i.e., the best one-predictor model; xR^2 = 0.140; Figure 3.1) and total RTOP score and room size (i.e., the best two-predictor model; xR^2 = 0.2043; Figure 3.2). Models with more than two predictors are not further discussed, as additional variables contributed minimally to the cross-validated R^2.

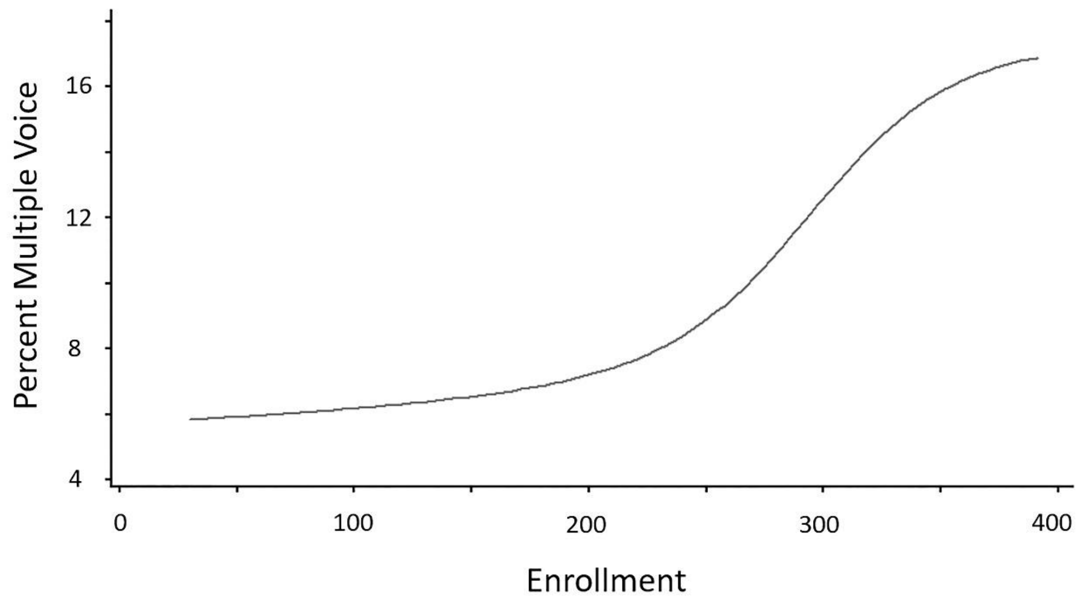


Figure 3.1. Two-dimensional fit response curve from NPMR, modelling section enrollment as a single predictor of percent Multiple Voice. Enrollment was the single variable in the best one-predictor NPMR model.

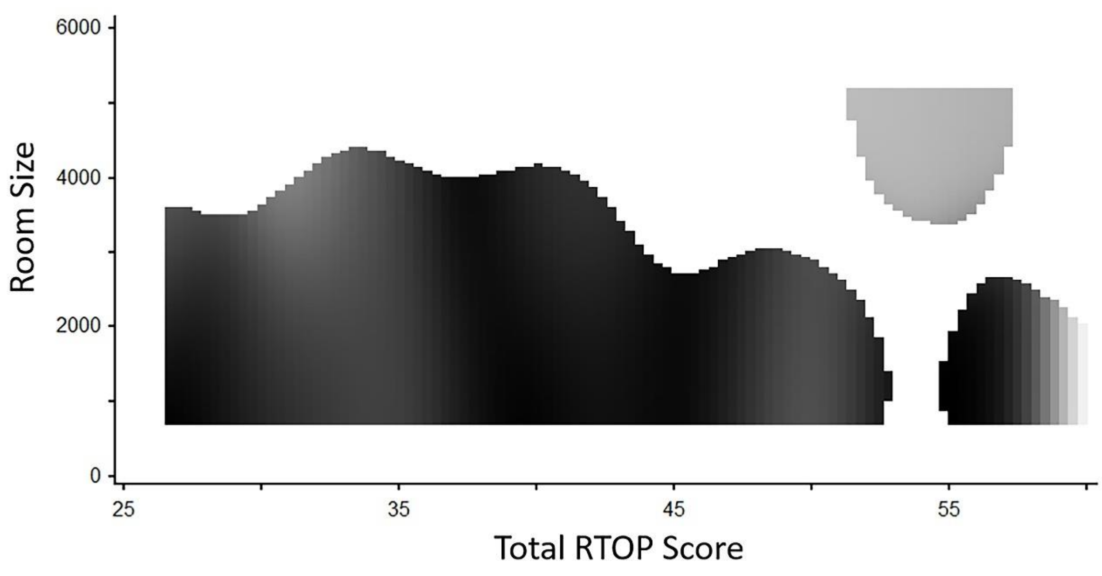


Figure 3.2. Three-dimensional contour response curve modelling total RTOP score and room size as the strongest predictors of percent Multiple Voice in a two-predictor model. Room size (square footage) and RTOP were the two variables in the best two-predictor NMPR model. The lightest colors represent the highest percent Multiple Voice detected by DART, grading into the darkest black that corresponds to the lowest percent Multiple Voice. Pure white represents values for which we had no data.

Total RTOP, enrollment, and room size were all significantly correlated, and room size and enrollment were the two most highly correlated predictors in our study (*r* = 0.974), thus effectively representing an equal measure of class size and capacity (Table 3.4). We found that the highest percentages of Multiple Voice were recorded in both: a) small classrooms taught by instructors with our highest values of total RTOP scores; and b) large classrooms taught by instructors with moderate to high total RTOP scores (Figure 3.2). The best one-predictor model, where we forced the single predictor to be total RTOP score, explained little variance in Multiple Voice (xR^2 = 0.0234).

Table 3.4. Pearson correlations between the best predictor variables in our models.

	% Multiple Voice (DART)	Total RTOP score	Room size (sq ft)	Enrollment
Total RTOP score	0.315*			
Room size (sq ft)	0.440**	0.381*		
Enrollment	0.460**	0.390*	0.974**	

Note: (*) denotes an alpha of 0.05 or less, while (**) denotes an alpha of 0.01 or less.

Discussion

Misalignment of the Decibel Analysis for Research in Teaching Tool

There could be multiple reasons why DART did not align well with an established measure of learner-centeredness (i.e., RTOP), and often underestimated the level of learner-centeredness for instructors scoring higher on the RTOP in our sample. Perhaps the singular focus of DART on sound within a classroom versus the more integrated focus of RTOP on both audio and visual observations within a classroom, caused misalignment in the output between these two instruments. Potentially DART captures different aspects of learner-centeredness than measured by RTOP, a phenomenon reported elsewhere for other instruments (Heim & Holt, 2018). Owens et al. (2018) even suggest that while DART may be a good indicator of general learner-centeredness, future work could investigate alignment of DART with other observation rubrics (e.g., Smith et al., 2013; Durham et al., 2018).

Additionally, technical aspects of our recording protocol likely affected the results. The use of lapel microphones by the instructors in our study may have interfered with how effectively student discussion in the classroom was detected by the audio recording devices, and represent a limitation of our study. If the microphones were mainly recording the instructor's voice because of their proximity to the instructor, this may explain why variance in percent Multiple Voice was fairly low (min = 0%, max = 30.8%). While this low variance was a limitation in our study, it also suggests a possible limitation in using DART among practitioners. Others have also found that to accurately capture students' voices in a classroom, multiple audio recording devices need to be set up throughout the room as to avoid singly capturing the instructor's voice simply due to proximity (Su, Dzodzo, Wu, Liu, & Meng, 2019). The positioning of audio recording devices in the classroom appears to be important for DART to collect sound accurately, yet further work is needed to clarify the optimal type of recording device and/or the placement of that device for everyday use by practitioners.

As there is a need for instruments that accurately gauge learner-centeredness of classrooms—which can easily be implemented by the "common educator"—and a need for undergraduate biology classrooms to be more active (Woodin, Carter, & Fletcher, 2010), observation protocols may provide benefits over other learner-centered instruments in that they utilize a more objective vantage point to both quantify learner-centered instruction and provide meaningful feedback to practitioners (Amrein-Beardsley & Popp, 2012; Durham et al., 2018; Eddy, Converse, & Wenderoth, 2015; Heim & Holt, 2018; Pratt & Collins, 2000; Sawada et al., 2002; Smith et al., 2013). While we initially expected DART would provide an effective and novel solution to the problem of practitioners' need for an accurate, off-the-shelf measure of learner-centeredness, this was not the case in our study. Calibration activities could have potentially improved the accuracy of DART (K. Tanner, pers. comm.); however, best practices and research on necessary calibration tasks are not widely available, further complicating the accessibility of DART for practitioners.

Big, Large Enrollment Classes Confound the Signal of the Decibel Analysis for Research in Teaching Tool

We found that as enrollment increased, as the single best predictor, so did percent Multiple Voice categorized by DART (Figure 3.1). Ultimately, more students in a classroom lead to more noise, whether from discourse related to course content, side conversations, or more individuals moving about the classroom. This finding suggests that DART may be biased in detecting learner-centeredness across classes of variable enrollments. Our best two-predictor model including room size (Figure 3.2) further suggests that large classes may bias DART's estimation of learner-centeredness, particularly since physically larger classroom spaces often amplify noise (Bradley, 2005; Seep, Glosemeyer, Hulce, Linn, & Aytar, 2000). While large enrollment classes can offer learner-centered environments (Knight, Wise, & Southard, 2013; Zagallo, Meddleton, & Bolger, 2016), it is unclear if DART can untangle these two sources of sound.

Encouragingly, the contribution of RTOP in our best two-predictor model was a near 50% increase over the variance explained in the one-predictor model by enrollment alone. While the overall variance explained by these two predictors was low, the addition of RTOP as a secondary predictor and its interaction with enrollment indicates that DART's prediction of learner-centeredness, at least minimally, aligns with another objective measure of learner-centeredness. Unfortunately, total RTOP score alone was not a good predictor of percent Multiple Voice (xR^2 = 0.0234). Although total RTOP scores had moderately low variance in our dataset, we argue that there was sufficient variance for our study (coefficient of variance = 24.67) to detect differences. While Bernstein (2018) suggests that DART could be a helpful tool in quantifying active learning in a classroom if further validated, many have found that observation protocols continue to provide the most accurate measurements of learner-centeredness in classrooms (Amrein-Beardsley & Popp, 2012; Durham et al., 2018; Eddy, Converse, & Wenderoth., 2015; Heim & Holt, 2018; Pratt & Collins, 2000; Sawada et al., 2002; Smith et al., 2013). Overall, DART's minor and interactive role in predicting learner-centeredness, and its misalignment with hand annotations in our study, weakens hope that it could be the panacea tool for practitioners.

Many Classroom Characteristics May Not Interfere with the Signal of the Decibel Analysis for Research in Teaching Tool

We included classroom characteristics in our models because we felt that some of these factors may unnecessarily distract from a signal of learner-centeredness. While enrollment and room size are clearly confounding factors when using DART, no other physical attributes of a classroom nor demographic factors were selected in the best models, which suggests that they were not contributing as much to classroom noise as we originally predicted.

Limitations of our Sample

We were mindful of pseudoreplication in our study, but neither instructor nor section identifiers were top predictors of percent Multiple Voice, thus this inherent redundancy was clearly not driving the overarching patterns we noticed in our models. Nine instructors teaching thirteen course sections were included in our sample to ensure consistency in course content being covered. However, greater variance in the classroom characteristic and demographic predictors, which could potentially be attained by increasing the number of course sections, instructors, and students sampled, could improve the fit of the models and allow us to measure which variables were most predictive of percent Multiple Voice with greater accuracy.

Conclusions

We found that enrollment was the best single predictor of percent Multiple Voice in a non-majors college biology course, and that total RTOP score and room size weakly predicted percent Multiple Voice when combined multiplicatively with one another. Specifically in regard to our research questions, we found that (1) DART did not align well with an established measure of learner-centeredness (i.e., RTOP), and often underestimated the level of learner-centeredness for instructors scoring higher on the RTOP in our sample, and that (2) only certain external variables (i.e., enrollment and room size) predicted DART output. We suggest that additional research is needed to clarify the types and positioning of audio recording devices necessary for effective DART analysis. Finally, RTOP and DART may be measuring distinct aspects of learner-centeredness, so the inclusion of other measures of learner-centeredness will be important to employ in future iterations of this research to determine whether DART is generally aligned with other instruments of learner-centeredness.

CHAPTER IV

INFORMAL LEARNING EXPERIENCES AMONG COLLEGE-AGE ADULTS: A PSYCHOMETRIC APPROACH AND APPLICATION OF THE INFORMAL LEARNING EXPERIENCES SURVEY

Contributions of Authors and Co-Authors

Manuscript in Chapter IV

Author: Ashley B. Heim

Contributions: Conceived study topic and design. Collected participant data. Organized and analyzed data. Wrote first draft of the manuscript.

Author: Emily A. Holt

Contributions: Provided feedback on methodology, analyses, and earlier versions of draft.

Abstract

While the autonomous nature of free-choice learning can have numerous positive effects on student learning in science fields, there is a lack of research on how college-age adults learn in informal learning settings. The purpose of this study was to quantitatively describe college-age adults' experiences at informal learning settings by developing and administering the novel Informal Learning Experiences Survey (ILES). We were interested in describing both the psychometric properties of the ILES as well as a practical application of the ILES using a sample population. We used psychometric analyses to test the reliability and validity of the ILES. We then used the full ILES with introductory biology undergraduates to describe the informal learning experiences in which college-age adults engage, and identified which factors best predicted frequency and number of types of informal learning experiences using linear hierarchical regression. We hope the ILES will (a) inform program directors at informal learning settings about how to better incorporate experiences designed for college-age adults, and (b) allow instructors of introductory college biology courses to reflect on and describe the backgrounds, prior experiences, and interests of their students related to learning in informal settings.

Introduction

What is Free-Choice Learning in Informal Learning Settings?

The National Science Teachers Association broadly describes informal learning environments in science as those that occur in out-of-school-time settings (NRC, 2009), including museums, science centers, zoos, and aquariums (MCZAs). Free-choice learning—or learning in which people choose what they want to learn about and for how long—in MCZAs both motivates students to persist in the sciences, and increases their understanding of science outside the formal classroom (NRC, 1996). At the K-12 level, free-choice learning is associated with increased student ownership of learning (Gardner 1991), increased understanding of science concepts, and increased persistence in the sciences (Adams & Branco, 2017; Drissner, Haase, Wittig, & Hille, 2014; Fadigan & Hammrich, 2004; Martell, 2008; Schwan, Grajal, & Lewalter, 2014; Subramaniam, 2002; Zimmerman & McClain, 2015). Informal learning experiences also benefit the learning of middle-aged and older adults (Alsop & Watts, 1997; Evans et al., 2005; Sachatello-Sawyer & Fellenz, 2000; Sachatello-Sawyer et al., 2002; Schwan et al., 2014). Learning at informal learning settings among college-age adults is relatively understudied. The majority of research on this age group has focused on the influence of social media on self-regulated learning (e.g., Dabbagh & Kitsantas, 2012; Kassens-Noor, 2012; Madge, Meek, Wellens, & Hooley, 2009) and the preparation of K-12 science teachers (Olson, Cox-Petersen, & McComas, 2001).

Theoretical Framework

While our study was exploratory and inductive by nature, our work leveraged Falk & Dierking's Contextual Model of Learning (2000), which describes a multi-factor framework for learning in informal learning settings based on personal (e.g., motivation, prior experience), sociocultural (e.g., social mediation), and physical contexts (e.g., visitor agendas, design of exhibits). All three of these components are integrated into the items on our Informal Learning Experiences Survey (ILES) and are broadly applicable to learning experiences across informal learning settings.

Personal context. Falk & Storksdieck (2005) describe the personal context of an informal learning experience as the personal history that a visitor brings into a learning situation, encompassing a visitor's (a) motives and expectations, (b) prior knowledge, experiences, and interest, and (c) autonomy to choose what to learn and for how long (p. 747). In our ILES, we describe and enumerate a person's reasons, or motives, for learning science in informal learning settings as well as their prior experiences at informal learning settings (i.e., as children or teenagers) within the personal context. The latter has been cited as a key factor influencing adults' decision to participate in informal learning opportunities (Falk & Needham, 2013). Pintrich and DeGroot (1990) explained that people are more likely to participate in learning experiences if they associate positive feelings and values with these experiences. Not only does prior interest influence a visitor's experience at an informal learning setting (Adelman et al., 2001; Adelman, Falk, & James, 2000; Csikszentmihalyi & Hermanson, 1995; Falk & Adelman, 2003), but so do less tangible aspects such as nostalgia (Borg & Mayo, 2005).

Sociocultural context. The sociocultural context is the influence of a visitor's social and cultural relationships on a learning scenario, encompassing a visitor's (a) within-group social interactions, and (b) outside-of-group social interactions (Falk & Storksdieck, 2005, p. 747). Our ILES gathers data on the sociocultural context by describing and enumerating the people with whom visitors usually engage at informal learning settings. Interactions with family members have been found to improve learning gains and scientific literacy for visitors of all ages in settings like museums, science centers, and zoos (Borun, Chambers, Dritsas, & Johnson, 1997; Crowley & Callanan, 1998). Often, family members facilitate learning in such settings by acquiring information from exhibits and discussing this information with others in their social group (Ellenbogen, Luke, & Dierking, 2004; Hilke & Balling, 1985; Naqvi, Venugopal, Falk, & Dierking, 1991). Beyond family members, visitor interactions with other visitor groups, volunteers, or staff can also influence the trajectory and quality of one's informal learning experience (Koran, Koran, Foster, & Dierking, 1988; Wolins, Jensen, & Ulzheimer, 1992).

Physical context. Lastly, the physical context incorporates any physical aspects within an informal learning setting that may contribute to how a visitor gains and applies knowledge. Collectively, these aspects may include: (a) visitor agendas, (b) orientation in the physical setting, (c) architectural design of the environment, (d) exhibit design and program development, and (e) reinforcing learning events that take place outside of the informal learning setting after the initial experience (Falk & Storksdieck, 2005, p. 747). Much of the physical context described above addresses elements of the environment when the participant is already on site, and we know anecdotally and from prior literature that college-age adults infrequently attend places of informal learning (Falk & Needham, 2013; Schwan et al., 2014). Thus, we focused on barriers college-age adults encounter in attempting to visit these settings, rather than physical characteristics experienced at the informal learning setting.

Our ILES captures some information regarding the physical context as people's barriers to visit informal learning settings. For adult visitors of lower socioeconomic status (SES), opportunities to visit MCZAs are often limited (Falk & Needham, 2013; Schwan et al., 2014). Zimmerman and McClain (2015) called attention to this SES bias in informal education research, emphasizing that MCZAs may cater more towards an educated and high SES audience, who can afford entry, rather than groups such as college-age adults who are often financially unstable or unable to procure transportation to MCZAs. Beyond financial barriers, we also evaluated if college-age adults' responsibilities interfered with their participation in informal learning environments.

Broader Impacts

Through our research, we aim to broadly describe the experiences of college-age adults at informal learning settings. Considering the alarming decrease in undergraduates persisting in science (Chen & Soldner, 2013), one solution may be to engage more college-age adults in informal learning experiences. Increased participation of undergraduates in learning opportunities at informal learning settings has the potential to improve students' content appreciation in formal learning environments (Wentzel & Brophy, 2014) and boost intrinsic motivation. Further, many college-age adults' future career skills will be learned informally; thus, free-choice learning experiences may better prepare them for a life as self-regulated learners (Zimmerman, 2002).

Development and administration of the ILES is a first step in addressing the knowledge gap of how experiences at informal learning settings influence the learning of college-age adults. We hope that our findings from the current study will encourage college faculty to implement more informal learning experiences in their curricula, or to consider the informal learning backgrounds, experiences, and interests of students via administration of the ILES. Additionally, we hope that program directors at informal learning settings might use the ILES to develop learning programs specifically for college-age adults.

Research Goals and Objectives

The purpose of this study was to quantitatively gain a better understanding of college-age adults' experiences at informal learning settings using the Informal Learning Experiences Survey (ILES). Our first research question was:

Q4.1 What do psychometric analyses suggest about the reliability and validity of the ILES?

Then as a first application of the ILES, we asked:

Q4.2 Among college-age adults, what/who are the most frequent (a) reasons and (b) barriers for learning science at informal learning settings; (c) people with whom college-age adults visit informal learning settings; and (d) informal learning settings visited as children/teenagers?

Q4.3 Which factors (a-d listed in Q4.2) best predict the frequency and number of types of informal learning opportunities in which college-age adults engage, including demographic characteristics?

For clarity, we first report common methods shared between both the psychometric (Q4.1) and application portions (Q4.2, Q4.3) of our research, followed by separate methods, results, and discussions for each.

General Methods

Ethics Statement

The Institutional Review Board of the University of Northern Colorado approved the procedures for this study (IRB #1227292-2; Appendix A2). Written informed consent was obtained by all participating students at the beginning of the study.

Site Description

All data were collected at a single, public four-year university in the western United States with an enrollment of nearly 9,000 undergraduates and 2,500 graduate students. Within this student population, approximately 59% of the undergraduates were white, 16% were Hispanic, and 4% were African Americans. Almost 85% of undergraduates were classified as in-state, and 34% of undergraduates identified as first-generation students. Nearly 64% of all undergraduates enrolled at this institution were females, while 36% were males.

Participants

We used a non-experimental research design and observed a single sample of a college-age adult population. Since college-age adults outside of academia are difficult to recruit, we narrowed our selection of participants to matriculating first- and second-year undergraduates within a biology major. We were interested in exploring informal learning experiences in the first half of students' college degree programs, because the first two years of a biology student's degree program are vital in retention in the biological sciences (Chen & Soldner, 2013).

Through convenience sampling, we sampled 453 students from five introductory 100-level biology courses, and complete survey responses from 441 students were analyzed. To improve response rates, students in all five of the participating courses were offered extra credit for completing the online survey. While volunteer participation sometimes results in non-response bias, the completion rate of 95% was proximal to the accepted average noted in psychological studies (Baruch, 1999). Further, across all five courses, student enrollment totaled 624, and our response rate of nearly 71% was sufficient based on an a priori power analysis conducted using G*power (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009), which estimated a minimum sample size of 91 students via linear regression analyses using a mean R^2 effect size of ~ 0 to 0.20 in biology survey research (Brownell et al., 2012; Nakagawa & Cuthill, 2007), an alpha of 0.05, five independent variables in the model, and a power estimate of 0.95.

The courses from which students were recruited were designed for biology majors, and represented the first two courses in an introductory biology series (cellular-molecular course, *n* = 3 lecture sections; ecology-evolution course, *n* = 1 lecture section). These biology course sections were taught by three instructors over the Fall 2018 semester (i.e., one instructor taught two sections). In the four participating courses, student enrollment ranged from 39 to 245 students (mean = 156 students per course). We assumed that informal learning experiences of undergraduate students enrolled in these four courses would be representative of the average first- or second-year undergraduate biology student, and furthermore, that our results would be applicable to individuals of this population, given our subset (i.e. college-aged people enrolled in an introductory biology course for biology majors). Exclusion criteria were defined as students aged 17 and under to maintain the exempt status of this research and avoid accommodation of a vulnerable non-adult population.

Most student participants (80.3%) were women, while 17.2% were men, and 2.5% were transgender men or women, gender-queer or gender-nonconforming, or another gender identity. Nearly 73% of students were white. Most students (90% of total sample) were enrolled in the cellular-molecular course. The majority of students identified as one of two majors (44.8% in Nursing, and 24.2% in Biology with a Pre-Health emphasis). This sample consisted of mostly first-year students (75.8%), and 80.5% of students were either 18 or 19 years of age. A large portion of students (14.7%) identified as transfer students from different institutions. Students grew up in households with a variety of annual incomes; nearly 80.6% of students' mothers and 71.7% of students' fathers earned at least a high school degree. Nearly 97% of sampled students were single or in a relationship but never married, and 97.5% did not have children. Nearly 72% of students reported that they spent the majority of their childhood in the state where the institution was located.

Data Analyses

All data analyses described below were conducted using SPSS (IBM Corp., 2016).

Part 1. Psychometrics of the Informal Learning Experiences Survey

Psychometric-Specific Methods

Instrument characteristics. We assumed that a primary underlying construct of student responses on the ILES (Informal Learning Experiences Survey) would be opportunity and upbringing. The ILES is composed of five items each with "choose all that apply" (CATA) responses, and the opportunity to write-in an "other" response. The findings presented in this paper are based on the second version of the ILES. The first version of the ILES was distributed in Fall 2017 as a pilot study (n = 334 students from the same two introductory biology courses that participated in the current Fall 2018 study; Appendix B), which allowed us to refine items in the ILES via exploratory factor analyses, item reliability analyses, and think-aloud interviews with introductory biology students. A brief summary of psychometric analyses from Version 1 of the ILES is presented in Appendix C, otherwise the current paper reports exclusively on analysis and use of Version 2.

While we below report on exploratory factor analyses (EFA) and item reliability analyses among the Version 2 CATA responses within each item, we want to emphasize that the ILES was developed primarily for practitioners to describe and better understand college-age adults' learning experiences at informal learning settings rather than as an instrument strictly for research purposes. However, for those who would like to adapt or use items from the ILES in their own research studies, we have provided results and interpretations from our EFA and item reliability analyses from Version 2 of the ILES.

Exploratory factor analysis (Step 1). Pattern matrices were used to interpret the content of each factor among the CATA responses within each item; see Table 4.1 for a summary of descriptive parameters for each factor derived from exploratory factor analysis (EFA). As responses on each of the CATA items within the ILES were not intended to be dependent on one another, we did not run a whole-survey exploratory factor analysis. Instead, we present results from EFA for each individual item to check for strong collinearity and patterns among CATA responses within the same item.

Table 4.1. Descriptive parameters of factors for items based on exploratory factor analysis.

Item	Item Description/ Name	Factors	Factor Description/ Name	CATA Responses included in each Factor	Mean Factor Scores	Standard Deviation	Reliability Estimate	CATA responses removed during EFA – Step 1	Overall item α after EFA – Step 1	CATA responses removed during item reliability analyses – Step 2	Final item α and improvement after item reliability analyses – Step 2[Δ]
1	Frequency/type of informal learning	1	General informal learning settings	1-9	0.26	0.156	0.860	None	0.852	8, 10, 12	0.894 (0.042)
		2	Outdoor learning/high entertainment value*	9-12	0.69	2.874	0.632				
2	Reasons for learning about science	1	Social/cultural & out-of-school time reasons	1, 2, 4, 6, 7	0.37	0.215	0.477	3, 10-12	0.484	8, 9	0.493 (0.009)
		2	Formal learning reasons	5, 8, 9	0.06	0.059	0.383				
3	Barriers	1	Personal responsibilities	3-6	0.40	0.212	0.532	11	0.380	2, 7, 8	0.465 (0.085)
		2	Limited resources	1, 2	0.60	0.229	0.207				
		3	Unique experiences	8, 9	0.15	0.123	0.222				
		4	Lack of interest/motivation	7, 10	0.04	0.033	0.113				
4	People†	1	Immediate family	1, 2	0.63	0.234	0.564	3, 6-9	0.272	5	0.316 (0.044)
		2	Extended family or children/unclear	4, 9	0.10	0.091	0.162				
5	Informal learning as children/teens	1	Common settings visited as children	2, 4-8	0.72	0.177	0.669	12	0.689	None	0.689 (0.000)
		2	Outdoor learning/nostalgic*	9-11	0.48	0.221	0.481				
		3	High entertainment value	1, 3, 9	0.87	0.103	0.337				

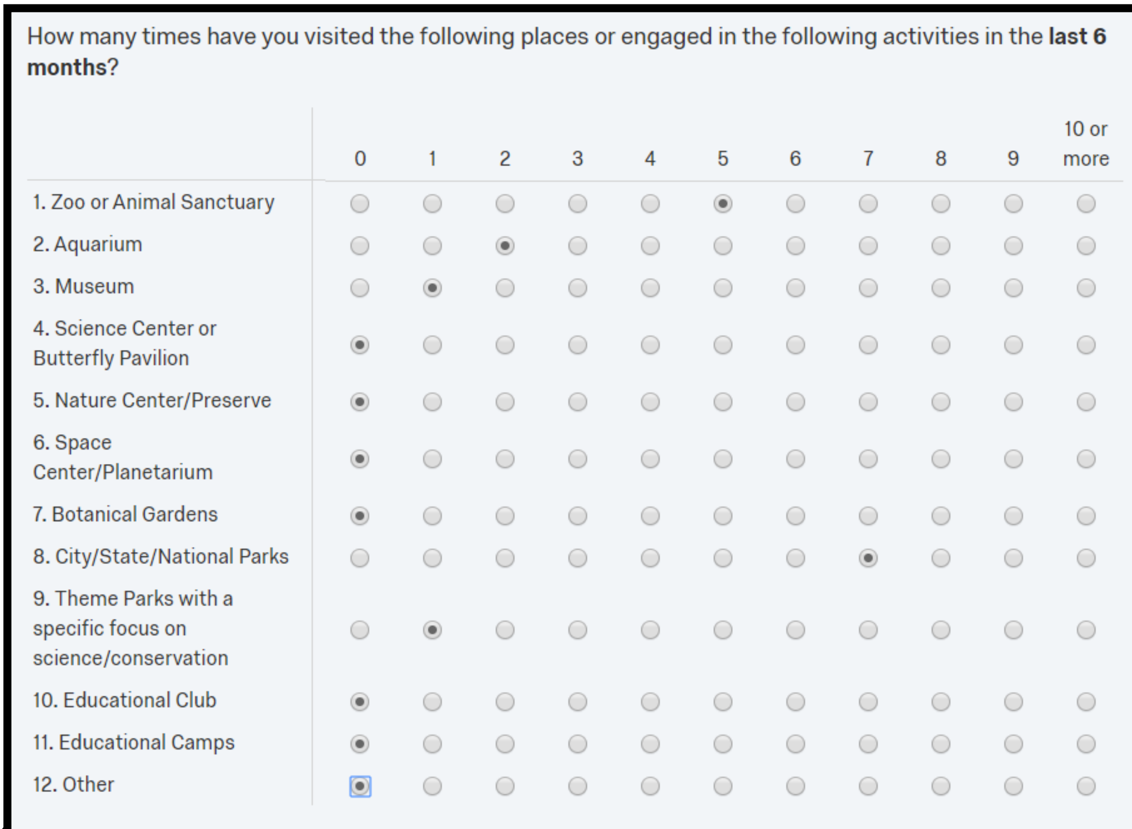
(*) indicates factors for which ambiguous wording of the CATA responses may have also contributed to factor loading.

(†) As significant others was the only CATA response that loaded onto the significant others factor, it was maintained in the item but could not be run in reliability analyses; hence, no EFA data are reported here.

(Δ) Note: CATA responses that are not listed in the "CATA Responses included in each Factor" column were removed during EFA before item reliability analyses were conducted. Further, improvements in alpha are noted in parentheses in the "Final Item α" column. CATA responses removed during item reliability analyses are in addition to CATA responses removed during EFA.

Numerical responses to Item 1 (i.e., Frequency/type of informal learning) were comprised of 12 CATA responses. Scores were created based on frequency of visitation in the last six months (sum of all informal learning visits from zero up to 10+ visits, across 12 environments, ranging from 0-120) and types of informal learning institutions visited in the last six months (sum of all settings a student visited, ranging from 0-12) for each student (Figure 4.1). We were also interested in exploring students' reasons for participating and not participating in learning at informal learning settings; the remaining four items of the ILES asked students to reflect on their reasons for learning about science (Item 2; Reasons for learning about science), barriers against participating in learning at informal learning settings (Item 3; Barriers), with which people they tended to engage in learning at informal learning settings (Item 4; People), and which informal learning settings they visited as children or teenagers (Item 5; Informal learning as children/teens; Appendix D).

It should be noted that because all items were in a CATA format, students also had the option to not select any of the listed options, which may have contributed to non-response bias on certain items. To create scores for the latter four items, selected CATA responses were summed to calculate a score for each ILES item (i.e., 12 reasons for learning about science in Item 2; 11 listed barriers in Item 3; 9 people in Item 4; and 12 learning settings visited as children in Item 5). Thus, if a student selected 4 of the 12 reasons for learning about science in Item 2, they would receive a score of 4 for that particular item.



How many times have you visited the following places or engaged in the following activities in the **last 6 months**?

	0	1	2	3	4	5	6	7	8	9	10 or more
1. Zoo or Animal Sanctuary	○	○	○	○	○	●	○	○	○	○	○
2. Aquarium	○	○	●	○	○	○	○	○	○	○	○
3. Museum	○	●	○	○	○	○	○	○	○	○	○
4. Science Center or Butterfly Pavilion	●	○	○	○	○	○	○	○	○	○	○
5. Nature Center/Preserve	●	○	○	○	○	○	○	○	○	○	○
6. Space Center/Planetarium	●	○	○	○	○	○	○	○	○	○	○
7. Botanical Gardens	●	○	○	○	○	○	○	○	○	○	○
8. City/State/National Parks	○	○	○	○	○	○	○	●	○	○	○
9. Theme Parks with a specific focus on science/conservation	○	●	○	○	○	○	○	○	○	○	○
10. Educational Club	●	○	○	○	○	○	○	○	○	○	○
11. Educational Camps	●	○	○	○	○	○	○	○	○	○	○
12. Other	●	○	○	○	○	○	○	○	○	○	○

To calculate the **frequency** of informal learning settings visited for this student, we would add up the total visits across locations (5 + 2 + 1, etc.) to get a total **frequency** sum of 16.

To calculate the score for **types** of informal learning places visited, the student would receive a score of 1 for any place they visited at least once in the past six months (e.g., zoo or animal sanctuary, aquarium, museum, city/state/national parks, and theme parks) which are then added up to calculate a total **types** sum. Here, the student visited five different settings, so they would receive a **types** score of 5.

Figure 4.1. Schematic representing how frequency and types scores are calculated for Item 1. Participants had 12 options to choose from on this CATA item, as well as 11 levels of visitation frequency. Thus, scores were created based on frequency of visitation in the last six months (sum of all informal learning visits, across 12 environments, ranging from 0-120) and types of informal learning institutions visited in the last six months (sum of all settings a student visited, ranging from 0-12) for each student. Note: Though not shown in this schematic, there was also a fill-in-the-blank option for students that chose “Other,” so that they had an opportunity to further describe their responses.

Five separate exploratory factor analyses (EFA) were conducted for each item on the ILES to investigate the clustering of CATA responses within each item; principal components analysis (PCA) was used as the default extraction method. Factors were maintained based on examination of scree plots and if they had initial eigenvalues greater than one, indicating the maximum number of potentially interpretable factors (i.e., based on the Kaiser-Guttman criterion). As our EFAs estimated multiple factor solutions, we opted to use direct oblimin rotation (delta = 0) to observe potential correlations among factors. At this step, we determined salient loading as factors with values greater than 0.3. Meaningful factors were then named and described, and poorly defined factors and/or poorly behaving CATA responses within ILES items were eliminated during EFA (Step 1) prior to running item reliability analyses (Step 2).

While we report on meaningful factors within the ILES based on removal of ill-fitting CATA responses for psychometric purposes in this portion of the paper, we maintained all items within the ILES for the second portion of this study to (a) more fully describe a sample population of college-age adults' informal learning experiences, and (b) because we contend that removal of poorly-performing CATA responses only minimally increased item reliability. Essentially, we argue that the costs of failing to fully describe students' experiences outweigh the benefit of item removal based on psychometric analyses. Thus, all analyses conducted in the application portion of our study are based on retaining all ILES items.

Item reliability analyses (Step 2). After running EFAs for each item, we conducted reliability analyses for each of the five items within the ILES, as well as for each item following CATA response deletions made during Step 2. If CATA responses were removed during EFA (Step 1), these CATA responses were not included in either item reliability analyses (Step 2). We calculated Cronbach's alpha for each item to determine internal consistency of CATA responses within each item. CATA responses were removed during Step 2, if the first item reliability analyses in SPSS suggested that deletion of individual CATA responses improved the overall reliability of a given item, even if marginally (i.e., Cronbach's alpha improved if individual CATA responses were removed).

Psychometric-Specific Results

Exploratory factor analysis (Step 1). Within Item 1 (i.e., Frequency/type of informal learning), EFA suggested that most informal learning settings loaded onto two factors: one factor we have labeled "general informal learning settings" (including CATA responses 1-9; Table 4.1). The CATA responses that simultaneously (i.e., Theme parks; CATA response 9) or exclusively (i.e., Educational clubs, Educational camps, Other; CATA responses 10-12) loaded onto a second factor may have done so due to the ambiguous wording of these responses, because students could have interpreted the responses in numerous ways (e.g., Educational camps or clubs might mean different things to different participants; Table 4.1). Further, CATA responses that loaded onto the second factor of Item 1 had themes of outdoor learning and high entertainment value in common. No CATA responses were removed during EFA. Oblique rotation converged in 8 iterations, and the two primary factors explained 53.85% of the common variance (Factor 1: 44.82%; Factor 2: 9.026%).

Within Item 2 (i.e., Reasons for learning about science), EFA suggested that most reasons for learning about science loaded onto two factors: one defined by social and cultural reasons and autonomous learning outside the classroom (including CATA responses 1, 2, 4, 6, and 7), and one defined more by formal learning (including CATA responses 5, 8, and 9). CATA responses that loaded onto non-meaningful factors (i.e., I feel culturally and socially accepted at these places, Just for fun. I find the experience enjoyable; CATA responses 10 and 3) or no factors at all (i.e., I volunteer at one or more of these places: CATA response 11) were removed during EFA (Bandalos & Finney, 2018). Oblique rotation converged in 12 iterations. These two factors explained 28.02% of the common variance (Factor 1: 16.30%; Factor 2: 11.72%).

Within Item 3 (i.e., Barriers), EFA suggested that most barriers against informal learning loaded onto four factors: one defined by personal responsibilities (including CATA responses 3, 4, 5, and 6); one defined by limited resources (including CATA responses 1 and 2); one defined by unique experiences at these institutions (including CATA responses 8 and 9); and one defined by lack of interest or motivation (including CATA responses 7 and 10). Oblique rotation converged in 14 iterations. These four factors explained 48.1% of the common variance (Factor 1: 16.0%; Factor 2: 11.6%; Factor 3: 10.7; Factor 4: 9.8%). The strongest Pearson correlation among CATA responses was measured between CATA responses 5 (i.e., Family responsibilities) and 6 (i.e., Social responsibilities; $r = 0.293$), perhaps because these two concepts are often highly interrelated. CATA response 11 (i.e. Other) was removed during EFA.

Within Item 4 (i.e., People), CATA responses describing people with whom students engaged in informal learning mainly loaded one of three factors: one describing the immediate family of most unmarried young adults (including CATA responses 1 and 2); one describing significant others (including CATA response 5); and one describing other family members or children of these young adults (including CATA responses 4 and 9). As Significant others was the only CATA response that loaded onto the Significant others factor, it was maintained in the item but could not be run in reliability analyses. It should be noted that Teachers/mentors (CATA response 6) and I prefer to go by myself (CATA response 7) each loaded negatively onto two separate factors (in addition to the three described above), and were hence removed prior to further reliability analyses. Surprisingly, Friends (CATA response 3) loaded negatively onto a separate sixth factor and was removed during EFA. Oblique rotation converged in 11 iterations. These three factors explained 42.8% of the common variance (Factor 1: 17.5%; Factor 2: 13.2%; Factor 3: 12.0%).

Within Item 5 (i.e., Informal learning as children/teens), informal learning settings that students visited as children or teenagers mainly loaded onto three factors: one factor we labeled “common informal learning settings visited as children” (including CATA responses 2, 4-8). Theme parks, Educational clubs, and Educational camps (CATA responses 9-11) loaded onto a second factor, again perhaps due to ambiguous wording or because they had themes of outdoor learning and nostalgia in common, as mentioned for Item 1. Zoo or Animal sanctuary, Museum, and Theme parks (CATA responses 1, 3, and 9) all loaded onto a third factor defined by a high entertainment value at these informal learning settings, yet Museum did so very weakly (<0.4). Oblique rotation converged in 8 iterations. These three factors explained 43.2% of the common variance (Factor 1: 23.1%; Factor 2: 10.9%; Factor 3: 9.2%). CATA response 12 (Other) was removed during EFA.

Most inter-CATA response correlations within each ILES item were weak ($r < 0.100$), suggesting that multicollinearity was not a concern for later reliability analyses.

Item reliability analyses (Step 2). While there is no agreed upon "acceptable" value of Cronbach's alpha in the science education literature (Taber, 2018), alphas for ILES items in our study, after EFA but prior to additional removal of items (in Step 1), had a broad range of reliability estimates (Item 1, $\alpha = 0.852$; Item 2, $\alpha = 0.484$; Item 3, $\alpha = 0.380$; Item 4, $\alpha = 0.272$; Item 5, $\alpha = 0.689$; Table 4.1). We removed between zero to three CATA responses for each item to improve reliability during item reliability analyses (in Step 2). Though these removals resulted in slight statistical improvements, the difference in alphas prior and following these removals was less than 0.1 in all cases (Table 4.1). Thus, for practicality, all CATA responses were retained for the application portion of this paper. Although EFAs of each item on the ILES—with their multiple CATA responses—loaded into more than one factor, low reliability estimates for multiple factors of one item suggested that we use only one summed score for each item (Taber, 2018; Table 4.1).

Psychometric-Specific Discussion

Our goal for Part 1 of this work was to explore the validity and reliability of this new instrument by answering our first research question (Q4.1). While the ILES was developed primarily for describing and better understanding college-age adults' experiences at informal learning settings, we recognize the importance of evaluating the psychometrics of a novel instrument for research purposes. While Items 1 (i.e., Frequency/type of informal learning) and 5 (i.e., Informal learning as children/teens) had acceptable Cronbach's alphas (Taber, 2018) and required minimal removal of items to improve reliability, Items 2-4 (i.e., Reasons for learning about science, Barriers, and People, respectively) had lower reliability scores and required removal of more items to improve reliability (Table 4.1). We suggest that based on our psychometric analyses, items on the ILES are variably reliable and more suitable for descriptive analyses.

We emphasize that although certain CATA responses were removed during Part 1 (psychometric analyses), we maintained all CATA responses for Part 2 of the current study. We felt retention of all CATA responses was critical—despite suggested removal in EFA and item reliability analyses—because many CATA responses were data-rich and provided important insight into the informal learning experiences of our sample, and often, removal of CATA responses only marginally improved item reliability. For example, although City, State, and National Parks (CATA response 8) of Item 1 was removed during reliability analyses, it was the most commonly visited informal learning setting among our participants (n = 1871 total visits; 86% of all students noted that they had visited a park in the past six months). Additionally, while we removed Friends (CATA response 3) during EFA, it was a frequently selected option among participants (i.e., 83% of students selected this option when completing the ILES). Thus, all CATA responses across ILES were used in the application portion of our study (i.e. Part 2) to ensure a robust description of informal learning experiences among college-age adults.

Part 2. Application of the Informal Learning Experiences Survey

Application-Specific Methods

Coding of variables. Item 1 of the ILES, our dependent variable, asked students to select a numerical value from 0-10+ for 12 responses; frequency was the sum of CATA responses (0-10) at the 12 provided informal settings, while type was the sum of the unique informal settings visited at least once within 6 months (Figure 4.1). Scores for Items 2-5 (i.e., Reasons for learning about science, Barriers, People, and Informal learning as children/teens), our independent variables of interest, were calculated by summing the number of CATA responses for each item (Appendix D). For all demographic items (Appendix E), with the exception of Item 13 (i.e., zip codes were converted to binary codes: within-state and out-of-state locations), response options were categorical and therefore had to be dummy coded for inclusion in the regression models.

Data analyses. We ran descriptive statistics to summarize the student sample, examine distributions and frequencies of the data, and assess appropriateness of the data to be included in later regression models, as well as answer our second research question (Q4.2). Crosstabulation analyses were conducted to examine differences in ILES item responses across demographic characteristics; p-values from Pearson chi-square tests represented two-sided asymptotic significance, and a Bonferroni-adjusted alpha of 0.0036 per test was used to maintain an error rate of 0.05 across all demographic variables. No demographic differences were detected via crosstabulation analyses, and thus are not discussed below.

Hierarchical linear regressions were used to answer our third research question (Q4.3), with frequency of visits to informal learning settings (i.e., the “frequency” model), as well as number of different settings visited (i.e., the “types” model; different summaries of ILES Item 1; see Figure 4.1) acting as the dependent variables in two separate models. We included four variables of interest (i.e., Reasons for learning about science, Barriers, People who accompany one at informal learning settings, and Informal learning settings visited as children/teens; ILES Items 2-5) and 14 demographic variables acted as the independent variables in both models. The R^2 values for each linear regression model were examined, as were the p-values and F-test for the R^2. Assumptions of linear regression were met (i.e., linearity, homoscedasticity, and inclusion of all relevant variables in the model). Variables were entered in two steps, with demographic variables tested at step one, and the five scores from ILES Items 2-5 added at step two, for each of the two models.

Application-Specific Results

Describing college-age adults' responses on the Informal Learning Experiences Survey (Q4.2). The most commonly visited informal learning setting among our participants was City, State, and National Parks (n = 1871 total visits; 86% of all students noted that they had visited a park in the past six months). The mean number of different types of informal learning settings visited by our sample in the previous six months was 4.87 (SD = 2.78).

Students reported that in the last six months their main reasons for learning about science at informal learning settings (i.e. ILES Item 2) were For fun and enjoyment (n = 353; 80%), To gather with friends and family (n = 252; 57%), and To learn about something new (n = 195; 44%). The top reported barriers against engaging in learning at informal learning settings (i.e. ILES Item 3) were Limited finances (n = 312; 71%), School responsibilities (n = 284; 63%), and Lack of transportation (n = 214; 48%) as well as Job responsibilities (n = 214; 48%). Students overwhelmingly noted that the people with whom they most commonly visited informal learning settings (i.e. ILES Item 4) were Friends (n = 368; 83%), Parents (n = 282; 64%), and Siblings (n = 273; 62%). Lastly, a majority of students had visited Zoos (n = 426; 96%), Museums (n = 407; 92%), Aquariums (n = 390; 88%), City, State, and National Parks (n = 376; 85%), Science centers or Butterfly pavilions (n = 317; 72%), Theme parks (n = 314; 71%), Space centers or Planetariums (n = 257; 58%), Botanical gardens (n = 256; 58%), and Nature centers/preserves (n = 232; 53%), as children or young teenagers (Figure 4.2; ILES Item 5).

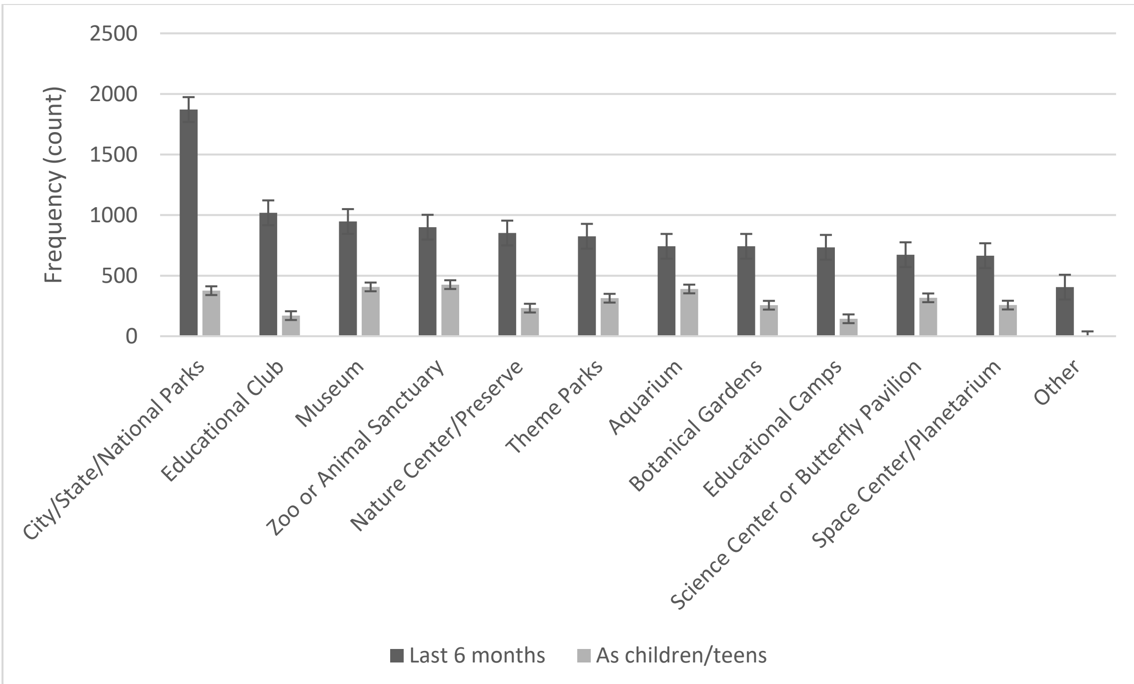


Figure 4.2. Most frequently visited informal learning settings among students in the last six months and as children/teenagers. City, states, and national parks were the most frequently visited places in the last six months. Further, students visited significantly more informal learning settings as young adults in the last six months compared to when they were children or teenagers. As scores for the “frequency” item were created by summing all of an individual’s informal learning visits (0 to 10+ visits for each setting, across 12 settings, total “frequency” scores for each student could range from 0-120); thus, this figure also incorporates multiple visits to the same location by individuals, which is why our findings are represented as frequency counts rather than percentage of students.

What predicts the frequency of informal learning opportunities in which college-age adults engaged (Q4.3)? The four ILES items, representing factors that encourage or prevent attendance to informal learning settings, simultaneously added at step 2 of the “frequency” model (*F* [4, 422] = 2.473, *p* = 0.001, R^2 = 0.095) improved the fit of the model beyond what was explained by the demographic variables in step 1 (*F* [14, 426] = 1.692, *p* = 0.055, R^2 = 0.053; Table 4.2), though neither model explained much variance in the frequency of visits. Two of the four variables of interest contributed uniquely to explaining a higher frequency of visits to informal learning settings: more Reasons for learning about science (*p* = 0.0001; *t* = 3.645; B = 1.671) and fewer Barriers against visiting informal learning settings (*p* = 0.001; *t* = -3.307; B = -1.536). Of the demographic variables, only higher estimated course grade contributed uniquely to explaining greater frequency of informal learning visits (*p* = 0.0001; *t* = 3.743; B = 2.369).

Table 4.2. Summary of Hierarchical Regression Analysis for Variables Predicting Frequency & Types of Informal Learning Opportunities in which Students Engaged.

		“Frequency” model				“Types” model			
Models/Steps		***B***	***SE B***	***β***	***p***	***B***	***SE B***	***β***	***p***
1. Demographic variables		18.170 (constant)	5.801 (constant)		0.055	7.435 (constant)	1.205 (constant)		0.038
14 Demographic Characteristics	**Race/Ethnicity**	-1.104	1.573	-0.037	0.483	-0.078	0.327	-0.012	0.811
	Gender	-0.110	0.983	-0.005	0.911	0.009	0.204	0.002	0.966
	Course	0.581	2.430	0.013	0.811	-0.327	0.505	-0.035	0.517
	Grade	2.233	0.639	0.176	0.001	0.358	0.133	0.136	0.007
	Major	-0.152	0.137	-0.057	0.270	-0.068	0.029	-0.123	0.017
	Year in School	-0.103	1.055	-0.006	0.923	-0.010	0.219	-0.003	0.964
	Transfer Student	-3.658	1.978	-0.108	0.065	-0.651	0.411	-0.093	0.114
	Age	-0.326	0.633	-0.041	0.607	0.053	0.131	0.032	0.684
	Income	-0.145	0.136	-0.053	0.287	-0.036	0.028	-0.063	0.206
	Mother Education	0.422	0.301	0.073	0.162	0.059	0.063	0.049	0.347
	Father Education	-0.192	0.270	-0.036	0.477	-0.024	0.056	-0.021	0.675
	Marital Status	-0.233	0.289	-0.037	0.441	0.044	0.060	0.035	0.464
	Number of Children	-0.203	2.665	-0.004	0.939	-0.767	0.554	-0.081	0.167
	Grew up in-state*	-1.714	1.493	-0.057	0.251	-0.622	0.310	-0.100	0.046

Table 4.2, Continued.

	Models/Steps	"Frequency" model *B*	*SE B*	*β*	*p*	"Types" model *B*	*SE B*	*β*	*p*
2. Independent variables of interest		16.100 (constant)	6.195 (constant)		0.001	6.817 (constant)	1.277 (constant)		0.0001
14 Demographic Variables	**Race/Ethnicity**	-0.944	1.555	-0.031	0.544	-0.057	0.321	-0.009	0.860
	Gender	-0.154	0.970	-0.008	0.874	0.010	0.200	0.002	0.961
	Course	-0.455	2.415	-0.010	0.851	-0.497	0.498	-0.053	0.319
	Grade	2.369	0.633	0.187	0.000**	0.379	0.131	0.144	0.004
	Major	-0.131	0.135	-0.049	0.331	-0.063	0.028	-0.113	0.024
	Year in School	0.065	1.048	0.004	0.951	-0.007	0.216	-0.002	0.973
	Transfer Student	-2.588	1.964	-0.077	0.188	-0.389	0.405	-0.055	0.337
	Age	-0.233	0.624	-0.029	0.709	0.092	0.129	0.055	0.474
	Income	-0.184	0.134	-0.067	0.173	-0.049	0.028	-0.086	0.078
	Mother Education	0.386	0.297	0.067	0.193	0.048	0.061	0.040	0.429
	Father Education	-0.272	0.269	-0.052	0.313	-0.041	0.056	-0.038	0.458
	Marital Status	-0.240	0.284	-0.040	0.398	0.037	0.059	0.029	0.533
	Number of Children	-0.316	2.631	-0.007	0.904	-0.868	0.543	-0.091	0.110
	Grew up in-state*	-1.223	1.486	-0.041	0.411	-0.523	0.306	-0.084	0.089
Variables of interest	**Reasons for Learning**	1.671	0.459	.211	0.000**	0.298	0.095	0.181	0.002**
	Barriers	-1.536	0.464	-0.179	0.001**	-0.408	0.096	-0.229	0.000**
	People	-0.543	0.606	-0.047	0.370	-0.050	0.125	-0.021	0.692
	Prior Experiences	0.327	0.297	0.056	0.272	0.132	0.061	0.108	0.032

*Refers to students who grew up in the same state where the current institution is located.
**$p < .0036$ (Bonferroni-adjusted).

What predicts the number of types of informal learning opportunities in which college-age adults engaged (Q4.3)? The four ILES items simultaneously added at step 2 (F [4, 422] = 2.938, p = 0.0001, R^2 = 0.111) improved the fit of the "types" model beyond what was explained by the demographic variables in step 1 (F [14, 426] = 1.785, p = 0.038, R^2 = 0.055; Table 4.1). Similar to the "frequency" model, two of the four variables of interest contributed uniquely to explaining different types of informal settings visited: more Reasons for learning about science (p = 0.002; t = 3.152; B = 0.298) and fewer Barriers against visiting informal learning settings (p = 0.0001; t = -4.265; B =

-0.408). However, no demographic variables contributed uniquely to explaining the types of informal learning settings visited in this second model.

Application-Specific Discussion

Reasons for learning about science. More reasons for learning about science predicted both how often college-age adults engaged in learning at informal learning settings and the diversity of settings they visited. Principal reasons reported for learning about science were: (1) for fun and enjoyment, and (2) to gather with friends and family. Interestingly, individuals also reported that they most often participate in free-choice learning at these settings with friends, perhaps reflecting the social nature of learning experiences at informal learning settings for college-age adults. Likewise, Falk and Gillespie (2009) suggested that the unique experiences offered through informal learning exhibits, and the emotions elicited by such experiences, may in part be due to the sociality often associated with visiting informal learning institutions. Further, Falk, Scott, Dierking, Rennie, & Jones (2004) found that interactive exhibits improved how students socially engaged in science learning.

The fun and enjoyment that individuals in our study associated with learning science at informal learning settings may be rooted in Pugh's (2004) idea of transformative experiences, in which students use science concepts for meaning making in their everyday lives and often become more motivated to learn science autonomously (Pugh, Linnenbrink-Garcia, Koskey, Stewart, & Manzey, 2010). Studies have also reported the appeal of autonomous learning among participants of informal learning opportunities, focusing on the notion that people are more willing to learn voluntarily about a topic when it directly relates to their daily lives (Alsop & Watts, 1997). Falk and

Dierking (2002) also emphasized that even in informal learning settings characterized more by entertainment than education, learning can still be a significant by-product of free-choice, environmentally-oriented experiences. Additionally, learning in outdoor or natural environments and direct encounters with nature can motivate people to learn about and become more aware of the natural world while simultaneously providing opportunities for leisure (Kellert, 1997; Kola-Olusanya, 2005; Negra & Manning, 1997).

Barriers against participating in informal learning. Fewer barriers among college-age students also contributed to more frequent visits to informal learning settings and a greater diversity of informal settings visited. The most frequently reported barriers against engaging in learning at informal learning settings within our sample were limited resources and other obligations. Our findings support previous reports that limited opportunities for visiting informal learning settings often exist due to one's socioeconomic status (SES) and lack of resources (e.g., financial, transportation, time; Falk & Needham, 2013; Schwan et al., 2014). However, this confirmation of SES bias associated with engagement in learning at informal learning settings reinforces the urgency to provide better learning opportunities for college-age adults who may not have the resources to participate in such activities outside a classroom environment. Additionally, if unique and engaging experiences are not available for certain age groups at informal learning settings, or visitors are not made aware of potential learning experiences and special events at informal learning settings, they are unlikely to allocate time to visit such places (Kola-Olusanya, 2005).

Estimated course grade. Our finding that higher estimated course grade was a predictor of higher frequency of engagement in informal learning settings supports what has been reported in the literature at the K-12 level, where learning in informal learning settings is associated with academic performance in the formal classroom (e.g., Arya & Maul, 2012; Barker & Ansorge, 2007; Drissner et al., 2014; Mayo, 2009; Subramaniam, 2002). However, while we offer a novel perspective and suggest that students who anticipate high performance tend to have higher visitation rates to places of informal learning, prior literature inversely suggests that visitation to informal learning settings predicts academic performance. Drissner et al. (2014) found that secondary school students who participated in an educational program at a botanical garden demonstrated more biological understanding and fewer biological misconceptions than their peers that did not participate. Many others have also found that engagement in free-choice learning programs improves student performance on classroom assessments and STEM-based achievement tests (Arya & Maul, 2012; Barker & Ansorge, 2007; Mayo, 2009; Subramaniam, 2002). In our study, students may have aligned learning in informal learning settings with their estimated course grades in the formal biology classroom, but we do not necessarily know that higher visitation rates caused students to have higher course grades. Additionally, as the ILES was administered to students approximately one month after the start of the semester, students were able to evaluate their academic performance based on course-based assessments and feedback.

General Summary & Conclusions

Limitations

As is true of most survey-based studies, ILES data are self-reported, which may result in bias (van de Mortel, 2008). Additionally, the sample used in this study was disproportionately comprised of females (80% of the total) and whites (73%); only sampled introductory biology students; and only sampled from one location in the intermountain west of the U.S. Thus, other more heterogeneous populations may respond differently on the ILES. Lastly, though students commented (via open-response survey feedback associated with Version 1) that they appreciated the CATA format of most items on the ILES, this format made data preparation and analyses challenging (e.g., non-response did not necessarily translate to missing items), hence the creation of total summed scores for each item.

The low reliability on certain items of the ILES (i.e., 2-4) should be interpreted with caution (Table 4.1), as the models using these items explained very little variance in the application portion of our study (i.e., Part 2). We recognize the low lack of fit within our models may indicate that other factors that we did not measure may have better explained the frequencies and number of types of informal learning experiences (e.g., other reasons for learning about science, barriers, and people that we may not have considered). However, the ILES was developed primarily for describing and better understanding students' free-choice learning experiences at informal learning settings rather than as a psychometrically-sound instrument for research purposes.

Practical Classroom Applications of the Informal Learning Experiences Survey

Our primary intention in developing the ILES was to provide a means for instructors or informal learning administrators to better understand and reflect on this population's experiences at informal learning settings. For college instructors or informal learning settings that intend to administer the ILES, we believe the most useful findings from completed surveys would be the percentage of individuals choosing each CATA response within each item. While the sums (i.e., scores) for each item can be used to broadly summarize the Frequencies/types of informal learning settings visited and the overall counts for each item (i.e., Reasons for learning, Barriers, People, and Informal learning as children/teenagers), these scores may not be as meaningful as identification of specific sites, reasons, and barriers.

Conclusions

Development and administration of the ILES is a first step in examining how experiences at informal learning settings influence the learning of college-age adults. Our findings could inspire faculty to consider the informal learning backgrounds, experiences, and interests of students via administration of the ILES. Additionally, we hope that program directors at informal learning settings might use the ILES to develop learning programs specifically for college-age adults, and college instructors may implement more informal learning experiences in their curricula. While certain items of the ILES had moderate to high reliability estimates (i.e., Frequency/variety of informal learning and Informal learning as children/teens) and could certainly be used for research purposes within biology and other STEM disciplines, the ILES in its entirety would presumably be best suited for reflective purposes (e.g., to better understand the learning experiences of undergraduates at informal learning settings in a biology course).

CHAPTER V

UNDERSTANDING BIOLOGY UNDER-GRADUATES' MOTIVATION AND INTEREST AFTER A LEARNING EXPERIENCE AT A REGIONAL ZOO

Contributions of Authors and Co-Authors

Manuscript in Chapter V

Author: Ashley B. Heim

Contributions: Conceived study topic and design. Collected participant data. Organized and analyzed data. Wrote first draft of the manuscript.

Author: Emily A. Holt

Contributions: Provided feedback on methodology, analyses, and earlier versions of draft.

Abstract

Free-choice learning, which often takes place in settings such as zoos, is where the learner has autonomy to choose what, where, how, and with whom to learn. As little is known about free-choice learning among undergraduates in informal settings and the potential of free-choice learning experiences at informal settings to engage more students in biology, we sought to answer the following research questions: (1) Does participation in structured versus free-choice learning experiences at the zoo relate to a biology student's motivation and interest to learn biology? (2) Does a biology student's status in their program (i.e., introductory or advanced) relate to baseline self-regulation, or a shift in motivation or interest after participating in a zoo trip? Students in both introductory and advanced biology courses were assigned to either a structured (i.e., structured agenda, led by chaperone) or free-choice (i.e., total autonomy) learning group during a visit to a regional zoo. Participating students completed a set of surveys before and after the zoo trip to gauge their incoming self-regulation and changes in motivation and interest to learn biology. We found that multiple aspects of motivation—including intrinsic motivation, career motivation, self-determination, and self-efficacy—increased after the zoo trip across all learning groups; however, the zoo trip benefit did not depend on how the trip was structured nor students' status as introductory or advanced.

Introduction

What is Free-Choice Learning?

Free-choice learning is where the learner has autonomy to choose what, where, how, and with whom to learn (Falk, Dierking, & Foutz, 2007). Generally, free-choice learning is also characterized by high intrinsic motivation of the learner to learn about the topic of their choice (Falk, Dierking, & Foutz, 2007). Often, free-choice learning takes place in informal learning (i.e., out-of-school-time) settings such as museums, science centers, zoos, and aquariums (MCZAs). Free-choice learning in MCZAs can motivate students to remain in STEM fields (Boekaerts & Minnaert, 1999; Falk & Storksdieck, 2010; Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000; Paris, 1997), increase their understanding of science beyond the formal classroom, and improve student engagement and sense of ownership in the classroom (Adams & Branco, 2017; Drissner, Haase, Wittig, & Hille, 2014; Fadigan & Hammrich, 2004; Schwan, Grajal, & Lewalter, 2014). Additionally, participation in free-choice learning experiences in informal settings has been linked to increased academic performance (Arya & Maul, 2012; Mayo, 2009) and greater conceptual understanding of biology (Drissner et al., 2014) among K-12 students in the formal classroom. As the majority of this research has focused on K-12 student populations, little is known about free-choice learning among undergraduates in informal learning settings, excepting for the preparation of K-12 science teachers in institutions such as museums (Olson, Cox-Petersen, & McComas, 2001).

Intrinsic Motivation during Free-Choice Learning Experiences

Intrinsic motivation is an individual's participation in an activity because he or she finds it personally rewarding and enjoyable (Sansone & Harackiewicz, 2000), or an individual's natural inclination to engage in a learning activity based on inherent interest (Ryan & Deci, 2000). The ability of an individual to construct personal meaning during a learning experience is often heavily aligned with his or her intrinsic motivation in that situation (Blumenfeld et al., 1991). The National Research Council even described the initial "learning" phase in informal science environments as experiencing interest, excitement, and motivation to learn about the natural world (NRC, 2009). Further, Falk, Dierking, & Foutz (2007) noted that free-choice learning generally encompasses one's intrinsic motivation to learn about a particular topic in an informal learning setting, though they also emphasized that not all learning in MCZAs would be intrinsically-driven (e.g., depending on the nature of the visit). Prior studies have found that visitors to informal settings were more intrinsically motivated to learn when able to develop their own agenda (Boekaerts & Minnaert, 1999; Falk & Storksdieck, 2010; Harackiewicz et al., 2000; Paris, 1997).

Extrinsic Motivation during Learning Experiences at Informal Settings

While informal settings such as MCZAs foster free-choice learning experiences among visitors, not every individual is intrinsically motivated to learn in such settings, particularly when the visit is required as part of a formal classroom curriculum. In these cases, students may be more extrinsically motivated by grades on an assignment associated with the visit; approval of the instructor or their peers; and accomplishing career goals (Paris, 1997; Sansone & Harackiewicz, 2000; Wentzel & Brophy, 2014). Free-choice learning in informal learning settings and learning in formal classroom environments can often be characterized by a greater reliance on intrinsic and extrinsic motivation, respectively (Csikszentmihalyi & Hermanson, 1995; Eshach, 2007; Lepper, Corpus, & Iyengar, 2005).

Self-Regulation during Free-Choice Learning Experiences

Self-determination is defined by the sense of control students have in learning a subject (Black & Deci, 2000), while self-efficacy describes students' personal beliefs that they can perform well in that subject (Lawson, Banks, & Logvin, 2007). Both aforementioned motivational aspects may influence a student's self-regulation, which Wigfield, Klauda, and Cambria (2011) describe as the means by which learners plan, monitor, and personally reflect on their performance to fulfill some sort of learning goal. Although formal classroom learning is more structured and compulsory free-choice learning (Wellington, 1990), exposing students to both structured and autonomous learning experiences in informal settings can increase student engagement and interest in the sciences as well as improve self-regulated learning skills (Bevan et al., 2010; Stuckey & Arkell, 2006). Prior studies have found that visitors who are more intrinsically interested in and motivated to learn a topic are more likely to develop learning goals for themselves at informal settings (Dierking, 2014; Dierking & Falk, 2009; Wilde, 2007).While little research has been conducted on undergraduates in informal learning settings, others have reported that adults set motivational goals for themselves at museums to more effectively plan their learning experiences (Falk & Storksdieck, 2010) and that undergraduates set learning goals for themselves during self-regulated learning activities in the formal classroom (Kitsantas & Dabbagh, 2011).

Theoretical Framework

Our research is theoretically founded on the Transformative Experiences Model (Garner, Kaplan, & Pugh, 2016; Garner, Pugh, & Kaplan, 2016). The Transformative Experiences Model, in the context of free-choice learning, describes how individuals can construct personal meaning from relevant concepts in their everyday lives. Specifically, such transformative experiences in the sciences are characterized by motivated use of a concept, expansion of one's perception, and the experiential value that an individual associates with a learning task (Kaplan, Sinai, & Flum, 2014; Pugh, Linnenbrink-Garcia, Koskey, Stewart, & Manzey, 2010). In our study, we believe that all of these factors may contribute to differential motivational outcomes among students participating in more autonomous versus more structured learning experiences at a zoo. For example, a student that has the opportunity to develop a personalized agenda during a zoo trip—rather than participate in a structured visit defined by limited choice—may be able to better adapt their learning experience based on inherent interest, value, and motivation to learn biology and achieve their personal learning goals. During free-choice learning, learners have the opportunity to make learning meaningful and interpret information in a way that is personally relevant to them (Ballantyne & Packer, 2006)—with the potential to engage in a unique transformative experience with continuing, rather than just short-term, effects (Rennie & Johnston, 2004).

Additionally, regarding academic level, advanced biology students may have more opportunity to engage in transformative experiences at informal settings compared to introductory students; advanced students often have higher levels of motivation to learn biology since they have had more experience in the discipline, more time to develop their interests, and thus more time to envision their biology-based career goals and possible selves (Pintrich & Garcia, 2012; Markus & Nurius, 1986).

Purpose, Research Questions, and Hypotheses

The purpose of this study was to describe the learning experiences of biology undergraduates at a zoo. Thus, we sought to answer the following research questions:

Q5.1 Does participation in structured versus free-choice learning experiences at the zoo relate to a biology student's motivation and interest to learn biology?

Q5.2 Does a biology student's status in their program (i.e., introductory or advanced) relate to baseline self-regulation, or a shift in motivation or interest after participating in a zoo trip?

We hypothesized that (1) student motivation or interest to learn about biology would increase after students participated in the free-choice zoo trip, and that (2) advanced biology students would generally have higher levels of motivation to learn biology since they have had more experience in the discipline, more time to develop their interests, and thus more time to envision their biology-based career goals and possible selves (Pintrich & Garcia, 2012; Markus & Nurius, 1986). Results will help us to understand whether exposure of undergraduates to more free-choice learning opportunities may mitigate the decreased persistence of students within biology, and may further improve individuals' intrinsic motivation, interest, and self-regulation to learn biology both within and beyond the formal classroom (Wentzel & Brophy, 2014; Zimmerman, 2002).

Methods

Ethics Statement

The procedures for this study were approved by the Institutional Review Board of the University of Northern Colorado (IRB #1301825-1; Appendix A3). Written informed consent was obtained by all participating students at the beginning of the study.

Participants

We conducted this observational study within one majors introductory (i.e., organismal biology, n = 39) and two majors advanced biology courses (i.e., animal behavior, n = 24, and mammalogy, n = 15) at a public 4-year university in the western United States. We used convenience sampling to select participants, and students were compensated with extra credit for participation. Since many students were simultaneously enrolled in both advanced biology courses included in our sample, students enrolled in both courses were advised to only participate in our study and be compensated in association with one of the courses.

While volunteer participation sometimes results in non-response bias, our total response rate of approximately 64% was proximal to the accepted average noted in psychological studies (Baruch, 1999). The number of participating students from which we received full pre- and post-survey responses from each group are noted parenthetically below.

Structure of the Regional Zoo Trip

Introductory biology course (n=33 students). Students enrolled in the introductory organismal biology course were randomly assigned to one of two required day-long zoo trips, offered the same weekend in September 2018. Students on the

Saturday zoo trip served as the “structured learning” group, and students on the Sunday zoo trip served as the “free-choice learning” group; however, students were not aware of the treatments to which they were assigned. The university provided students in both groups with free transportation and free admission to the zoo, to limit potential barriers to attendance. Further, though assigned to different treatments, students in both groups spent the same amount of time at the zoo each day (i.e., approximately seven hours, not including transit time).

Structured learning group (n=16 students). Students in the “structured learning” group were required to complete a structured assessment during their zoo visit, hereafter referred to as the Structured Zoo Content Assessment (Appendix F), which aligned with specific zoo exhibits and focused on topics such as taxonomy and adaptations. This assessment was provided at the start of the zoo trip, once students had entered the front admission gates, and was collected at the trip’s conclusion. As this assessment was developed for the structured learning treatment at the zoo rather than for a course assignment, students received credit for completing the handout rather than for correctness of responses. Students in the structured learning group were also given a visitor agenda, including a zoo map and timeline, which they were required to follow; this agenda described the exhibits students were expected to visit in a particular order, as well as the duration of time to spend at each exhibit.

To ensure that students adhered to the visitor agenda and had intentional, structured learning experiences, we further organized students on the Saturday zoo trip into three smaller groups each led by two graduate teaching assistant “chaperones” at the start of the day; students had no input regarding which peers composed each small group.

Each of these smaller groups was composed of six to seven students. Each small group had a unique visitor agenda to follow; while recommended durations and order of exhibits were similar among the three agendas, each group had a different starting location in the zoo to avoid overlapping of groups at the same exhibit. Each agenda also scheduled in two 20-minute zookeeper talks or demonstrations, though the topic of each of these talks or demonstrations differed among agendas due to limited daily showtimes. All groups had one hour scheduled for lunch, and thirty minutes at the end of the trip allocated for visiting the gift shop.

Free-choice learning group (n=17 students). Students in the "free-choice learning" group were required to complete a less structured, more general Free-Choice Zoo Content Assessment (Appendix G1) during their zoo visit, which did not necessary align with specific zoo exhibits and focused on broad topics such as taxonomy and organismal diversity. This assessment was provided at the start of the zoo trip, once students had entered the front admission gates, and was collected at the trip's conclusion. As in the structured group, students received credit for completing this Free-Choice Zoo Content Assessment.

Students in the free-choice learning group were given autonomy to choose the exhibits they wanted to visit, in whichever order, and for whatever duration they preferred. Therefore, students on the Sunday field trip were not assigned to smaller groups, were not supervised by chaperones, and did not have specific visitor agendas to follow. Students in the free-choice learning group, however, were required to track the order of exhibits they visited, including duration of time visited and any talks or demonstrations attended, on a blank map of the zoo. Students in this group did not have a scheduled time for lunch or to visit the gift shop, as individuals developed their own agendas.

Advanced biology courses (n=17 students). Students in the advanced biology courses in our study participated in the zoo trip on a Saturday in early October 2018. Again, students in these courses also received free admission, were offered free transportation, and spent approximately the same amount of time at the zoo as the introductory students.

Structured learning group (Animal Behavior; n=17 students). The instructor of the Animal Behavior course had a zoo trip required as part of the curriculum, including a structured ethogram assessment and animal behavior tours with zoo staff. Participating students from this course were identified as the structured learning group among the advanced biology students as they had limited autonomy in what they chose to do at the zoo. After participating in animal behavior tours led by zoo staff, students were able to explore the zoo individually or in groups to complete their ethogram assignments observing the animal species of their choice; most students spent the majority of post-tour time at the zoo completing these assignments. As this was a pre-determined component of the animal behavior course and not open to manipulation for our research, we define the advanced structured learning group as having more structure and less choice. In contrast, we define the advanced free-choice learning group, described below, as having a lack of structure and unlimited choice.

Free-choice learning group (Mammalogy; n=0 students). While we attempted to establish an advanced free-choice learning group, the sample size was small (n=3) and complete pre- and post-responses (i.e., matched data) were not received from any of the participating students. Thus, we could not include these data in our analyses and have an unbalanced design as a result. Similar to students in the introductory free-choice learning group, advanced biology students in this treatment had autonomy to choose the exhibits they wanted to visit, in whichever order, and for whatever duration they preferred. To ensure consistency among treatments, however, we did develop a general assessment similar to that for the free-choice introductory students (Appendix G2), which asked broad questions related to mammalogy and whether students chose to attend any zookeeper talks or demonstrations.

Assessments Administered Before and After the Zoo Trip

Pre-zoo trip assessments. One week prior to the scheduled zoo trips, all participating students were asked to complete four pre-zoo trip questionnaires online via Qualtrics. These questionnaires (described below) were intended to gauge students' motivation, self-regulation, and baseline interest in biology prior to visiting the zoo, as well as their prior experiences at zoos.

Prior Experiences at Zoos Questionnaire. To describe students' prior experiences at zoos and particularly the regional zoo used in this study, students were asked to complete a short questionnaire composed of four multiple-choice items that we created. All four items from this questionnaire are available in Appendix H, though we only used Items 1 and 3 in our analyses. As prior experience at a free-choice or informal learning setting may influence a visitor's learning on subsequent trips (Falk & Dierking, 2000), this questionnaire helped better describe the learning experiences of our student sample.

Learning Self-Regulation Questionnaire. We adapted the Learning Self-Regulation Questionnaire (LSRQ; Black & Deci, 2000; Ryan & Connell, 1989; Williams & Deci, 1996) to be relevant for biology students. We intended this metric to provide further insight into students' extrinsic motivations related to learning biology concepts at a zoo; sample items are available in Appendix I. This instrument was composed of twelve 7-point Likert-like scale items that characterized student responses on a spectrum from (a) controlled regulation (i.e., external or introjected regulation; $\alpha = 0.67$) to (b) autonomous regulation (i.e., identified or intrinsic regulation; $\alpha = 0.75$). While external regulation involves doing something for reasons completely external to oneself, introjected regulation is slightly more internalized and involves behaving in a certain way to feel worthy or avoid negative feelings (e.g., guilt)—often due to social pressures (Ryan & Deci, 2000). Identified regulation is further internalized motivation to do something, and involves the individual valuing a behavior and performing an action because they find it personally important or relevant (Ryan & Deci, 2000). Lastly, intrinsic regulation is closely aligned with intrinsic motivation (e.g., behaviors are aligned with self-values and ideals) but distinct in the sense that the individual is still not engaging in behaviors because of personal enjoyment (Ryan & Deci, 2000). This questionnaire—which has been identified as reliable and valid in the context of undergraduate science courses (Black & Deci, 2000)—was administered solely prior to the zoo trip to gain a better understanding of students' anticipated self-regulated learning during the zoo trip and analyze whether a difference existed in baseline self-regulation between introductory and advanced students.

Science Motivation Questionnaire-II. We adapted the Science Motivation Questionnaire-II (SMQ-II; Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011) for biology; this instrument was composed of 25 5-point Likert-like scale items that quantified how undergraduate students think and feel about their biology courses and about learning biology in general. Five motivational components were included within the SMQ-II: intrinsic motivation, career motivation, self-determination, self-efficacy, and grade motivation (Glynn et al., 2011).

We chose the SMQ-II over other motivation instruments such as the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich & DeGroot, 1990) because the aforementioned components of the SMQ-II have been shown to be valid and reliable within the context of both majors and non-majors undergraduate biology courses (Glynn et al., 2011). Others found that all five of the scales within the SMQ-II had moderate to high reliability estimates ($\alpha = 0.81$-0.92), while those within the MSLQ ranged from low to high ($\alpha = 0.52$-0.93; Pintrich & DeGroot, 1990).

Personal Interest in Biology Metric. We adapted the Personal Interest in Biology (PIB) measure from portions of the "Initial Interest" scale developed by Harackiewicz, Durik, Barron, Linnenbrink-Garcia, and Tauer (2008) and portions of the "Personal Interest" and "Meaningfulness" scales developed by Mitchell (1993) to better understand individual student's intrinsic interest in biology prior to and after the zoo trip. The aforementioned scales have been found to be both valid and reliable in undergraduate and high school courses ($\alpha = 0.90$, Harackiewicz et al., 2008; $\alpha = 0.77$-0.92, Mitchell, 1993). This instrument was composed of eight 5-point Likert-like scale items and was intended to assess students' personal interest in learning biology across a "value" scale and a “feeling” scale. While the value scale gauged how students perceived the practicality and usefulness of learning biology, the feeling scale aimed to measure students’ affect and emotions related to learning biology. All eight items are available in Appendix J.

Post-zoo trip assessments. Approximately two months after each zoo trip, originally recruited students were asked to complete a set of post-zoo trip questionnaires including the same items from the SMQ-II and PIB they completed in the pre-zoo trip questionnaires, as described above. We intended these post-zoo trip assessments to be a measure of whether student motivation and interest in biology changed after the zoo trip. We administered post-zoo trip surveys to students two months after the zoo trip rather than immediately after the zoo trip, as others have reported that short-term participation in free-choice learning experiences at informal settings often takes several days to weeks to have an impact on students (Bogner, 1998; Drissner et al., 2014; Rideout, 2005). In total, 89% of introductory biology students (33 of 37 participating students) and 71% of advanced biology students (17 of 24 participating students) completed both the pre- and post-zoo trip surveys; these are the only data we analyzed, thus no unmatched data are presented below.

Data Analyses

We ran eleven individual ANOVAs on student responses from the Prior Experiences at Zoos Questionnaire (2 items), Learning Self-Regulation Questionnaire (2 scales), Science Motivation Questionnaire-II (5 scales), and Personal Interest in Biology metric (2 scales) as response variables to characterize differences over time, by structure of the zoo trip and by academic level of the students. Post survey scores were used as the response variable for most models. We analyzed both the main effects and interactive effects of pre-responses with structure and level. Due to the lack of data from the advanced free-choice learning students in our study, ANOVAs that tested for pre-test and learning group (i.e., free-choice vs. advanced) interactions combined introductory and advanced students in the structured treatment, but included only introductory students in the free-choice learning group.

For data that we only collected prior to the zoo trip (i.e., Prior Experiences at Zoos Questionnaire and Learning Self-Regulation Questionnaire), we used the pre-survey score as the response variable and level (i.e., introductory or advanced) as a factor. As students had not yet participated in the zoo trip when they completed pre-surveys, we were not interested in comparing between structured and free-choice learning groups in these analyses. However, we did not find significant differences between structured and free-choice learning students regarding recency and frequency of zoo visits, nor regarding autonomous and controlled regulation (i.e., self-regulation), when adjusting for multiple comparisons. This suggests that there was no baseline differences across these four scales among our student sample. We used a Bonferroni-adjusted alpha of 0.0045 to account for these multiple comparisons. We used item reliability analyses via the "scale" function in SPSS to assess the internal consistency of items in each survey scale with our sample population; Cronbach's alpha was calculated for each scale. All quantitative data analyses were conducted using SPSS (IBM Corp., 2017).

Results

All scales of the four instruments used in this study were found to be moderately to highly reliable with our sample population (Table 5.1; α=0.630-0.925; Taber, 2018).

Table 5.1. Summary of ANOVA comparisons and reliability tests across survey scales.

Instrument[x]	Scale	Pre vs. Post	Structured vs. Free-choice	Introductory vs. Advanced	Reliability (α) Pre-test	Reliability (α) Post-test
PEZ	Recency of Zoo Visits*	N/A	N/A	- ($F_{1,48}$=0.354; p=0.555)	0.73	N/A
	Frequency of Zoo Visits*	N/A	N/A	- ($F_{1,48}$=2.616; p=0.902)		N/A
LSRQ	Autonomous	N/A	N/A	- ($F_{1,48}$=6.135; p=0.017)	0.70	N/A
	Controlled	N/A	N/A	- ($F_{1,48}$=1.621; p=0.209)	0.63	N/A
SMQ-II	Intrinsic Motivation	+ ($F_{10,29}$=4.171; p=0.001)	- ($F_{3,29}$=0.035; p=0.853)	- ($F_{5,29}$=0.015; p=0.902)	0.85	0.83
	Career Motivation	+ ($F_{8,28}$=16.738; p=0.0001)	- ($F_{7,28}$=1.149; p=0.293)	- ($F_{4,28}$=7.863; p=0.009)	0.79	0.89
	Self-Determination	+ ($F_{11,23}$=4.715; p=0.001)	- ($F_{6,23}$=1.141; p=0.297)	- ($F_{5,23}$=0.002; p=0.965)	0.82	0.82
	Self-Efficacy	+ ($F_{11,23}$=8.283; p=0.0001)	- ($F_{6,23}$=0.068; p=0.796)	- ($F_{6,23}$=2.248; p=0.147)	0.86	0.85
	Grade Motivation	+ ($F_{8,30}$=4.254; p=0.002)	- ($F_{4,30}$=2.025; p=0.165)	- ($F_{4,30}$=2.966; p=0.095)	0.74	0.80
PIB	Value	- ($F_{7,32}$=3.576; p=0.006)	- ($F_{5,32}$=0.026; p=0.873)	- ($F_{3,32}$=0.412; p=0.525)	0.82	0.81
	Feeling	+ ($F_{8,32}$=6.799; p=0.0001)	- ($F_{4,32}$=1.301; p=0.262)	- ($F_{2,32}$=0.544; p=0.466)	0.925	0.89

(-) indicates we found no difference between scores, based on an adjusted alpha of 0.005.
(+) indicates we found a significant difference between scores, based on an adjusted alpha of 0.005.
(*) Recency and frequency of zoo visits are items on the PEZ, not scales, thus why there is a single alpha reported for reliability.
([x]) Abbreviations of instruments: PEZ = Prior Experiences at Zoos Questionnaire; LSRQ = Learning Self-Regulation Questionnaire; SMQ-II = Science Motivation Questionnaire-II; PIB = Personal Interest in Biology Metric
Note: All treatment results are interactive effects with time, excepting for results of the PEZ and LSRQ scales

Prior Experiences at Zoos Questionnaire

Prior experiences at free-choice or informal learning settings can influence a visitor's learning on subsequent trips (Falk & Dierking, 2000). We found that 44.1% of all participating students reported visiting a zoo in the last year. We did not find any differences in recency of zoo visits between introductory and advanced biology students (p=0.56; Table 5.1). Further, 38.2% of all participants reported visiting zoos just once a year; 23.5% reported visiting zoos 2-3 times a year; and 38.2% reported that they never visited zoos. No significant differences were found in frequency of visits to zoos between introductory and advanced biology students (p=0.11; Table 5.1).

Learning Self-Regulation Questionnaire

While there was a trend that advanced students scored higher on the autonomous scale of the LSRQ ($p=0.017$; Table 5.1) compared to introductory students, this difference was not significant when adjusting for multiple comparisons. Further, we found no significant difference in scores on the controlled scale of the LSRQ between introductory and advanced students ($p=0.209$; Table 5.1). We did not compare autonomous and controlled scale scores between students in the structured and free-choice learning groups, as the LSRQ was administered before students participated in their treatment groups at the zoo.

Science Motivation Questionnaire-II

We calculated significant increases from pre- to post-scores for four scales of the SMQ-II (i.e., intrinsic motivation, $p=0.001$; career motivation, $p=0.002$; self-determination, $p<0.0001$; self-efficacy, $p=0.001$; Table 5.1) across all participants, but did not find that increases from pre- to post-scores were dependent upon the treatment groups (i.e., structured learning and free-choice learning, introductory and advanced students). We also found a significant decrease in pre- to post-scores across all participants on the grade motivation scale ($p=0.002$; Table 5.1), but again this reduction in grade motivation was not dependent upon treatment groups. Further, while there was a trend that advanced students scored higher on the career motivation scale of the SMQ-II compared to introductory students ($p=0.009$; Table 5.1), this difference was not significant at the adjusted Bonferroni alpha of 0.0045. Similarly, we found that self-efficacy scores were slightly higher in the structured group compared to the free-choice group, though this interaction with time was not significant ($p=0.034$; Table 5.1).

Personal Interest in Biology Metric

We calculated a significant increase ($p<0.0001$; Table 5.1) from pre- to post-scores for the feeling scale of the PIB across all participants, but did not find that increases from pre- to post-scores were dependent upon the treatment groups (i.e., structured learning and free-choice learning, introductory and advanced students). The value subscale scores of the PIB showed no differences between treatment groups or between time periods ($p=0.158$-0.827; Table 5.1).

Discussion

Prior to discussing our results, we want to emphasize that based on the lack of control groups in our study, we cannot state with absolute certainty that the zoo trip was what influenced changes in motivation, interest, and feelings related to learning biology on students' pre- to post-survey scores. While we did not have a comparable non-zoo trip "control" group due to most students across the three participating courses participating in the zoo trips, we did have one introductory student that was not able to participate in the scheduled zoo trip but did complete both the pre- and post-surveys. We qualitatively observed that this student's scores either remained the same or decreased on the post-survey across all scales of the SMQ-II and PIB metric. Though this pattern is strictly qualitative and based on one individual, thus not sufficient to make any definitive claims, it may provide insight into the idea that the improvements in motivation and positive feelings we observed across other participating students were due to zoo trip participation rather than content learned in the formal classroom between administration of the pre-and post-surveys. Thus, in our discussion below, we present our findings as differences we discovered between groups who participated in a zoo trip, yet we are cognizant that it could have been the zoo trip, the instruction in those intervening two months, or a mix of the two driving the changes in motivation, interest, and feelings related to learning biology that we observed. We also report what prior literature has found in the context of free-choice learning in informal settings.

No Difference if Zoo Trip is Structured Versus Free-Choice

Our primary and most interesting finding was that all students demonstrated improvements in various types of motivation, had more positive feelings about learning biology, and were less motivated by grades from pre- to post-surveys, regardless of whether they were assigned to the structured or free-choice group. Although the literature has historically concluded that free-choice learning is always more effective than structured learning (Drissner et al., 2014; Falk & Storksdieck, 2010; Schwan et al., 2014), in our study we found that the level of structure incorporated into a learning experience at the zoo does not matter. While others have reported that structured assessments and chaperones may limit the learning opportunities and interest of students visiting informal learning settings (Ballantyne, Fien, & Packer, 2001; Randol, 2004), students in our study that participated in a more structured learning experience at the zoo benefitted in multiple aspects of motivation just as much as students in the free-choice learning group.

The literature suggests that motivation measured by the SMQ-II is generally unchanged following formal learning experiences and only shifts with the introduction of informal experiences. For example, others have found that college STEM students’ motivation to learn decreases over a semester in a strictly formal classroom setting, using the SMQ-II (Wendel, Young, Esson, & Plank, 2016). Alternatively, Meesuk and Srisawasdi (2014) found that high school students in a chemistry course conveyed higher motivation and enjoyment to learn science after engaging in more free-choice educational computer games compared to their non-game playing peers, using the SMQ-II. Yamamura and Takehira (2017) also reported grade motivation on the SMQ-II tends to increase in the formal college science course over time, as students become more motivated to learn based on a desire to receive high grades. Additionally, Drissner et al. (2014) reported that secondary school students who engaged in a day-long free-choice learning experience in environmental science had more positive feelings related to learning biology than their peers who did not participate, at least in the short-term. Harackiewicz, Tibbetts, Canning, and Hyde (2014) noted that learning experiences which promote interest among individuals often lead to more positive feelings about that learning experience, which in turn can further increase interest in the subject matter.

No Difference if Students are Introductory or Advanced

While some students had visited a zoo in the last year—including the zoo where our study was conducted—nearly 40% of participating students noted that they generally did not visit zoos, and nearly 70% of students reported that they did not generally visit the zoo of interest. Not only do these findings from the Prior Experiences at Zoos Questionnaire emphasize the importance of better understanding the learning experiences of undergraduates in informal settings, but they are supported by anecdotal survey results collected by the authors regarding barriers to attending informal learning settings and others who described restricted access to free-choice learning based on limited resources

(Schwan et al., 2014). The limited prior experiences at zoos among undergraduates in our current study are unsurprising; others similarly find that opportunities at informal learning settings tend to be biased towards high socioeconomic status, educated adults (Falk & Needham, 2013; Zimmerman & McClain, 2015).

Interestingly, in the current study we also found that advanced students had not visited zoos more recently nor more frequently than introductory students; rather, they just have more formal exposure to biology topics through coursework, which did not manifest in higher initial motivation and interest levels as we had expected. Additionally, all participating students in our study had similar starting levels of motivation and interest—regardless of whether students identified as introductory or advanced. While there was a trend that advanced students scored higher on the autonomous scale of the LSRQ ($p=0.017$; Table 5.1) compared to introductory students, this difference was not significant when adjusting for multiple comparisons. While others have reported that advanced college students often have higher levels of motivation to learn since they have had more experience in the discipline, more time to develop their interests, and thus more time to envision their career goals and possible selves (Pintrich & Garcia, 2012; Markus & Nurius, 1986), our findings contradict this.

Relating our Findings to the Transformative Experiences Model

Our original theoretical hypothesis suggested that all of the components of a transformative experience in the sciences—motivated use of a concept, expansion of one's perception, and the experiential value that an individual associates with a learning task (Kaplan, Sinai, & Flum, 2014; Pugh et al., 2010)—may contribute to differential motivational outcomes among students participating in more autonomous versus more structured learning experiences at a zoo. However, we ultimately found that even if a student has limited autonomy to create their own personal agenda (i.e., our structured groups), there was no difference in motivation, interest, and positive feelings related to learning biology between students in the structured versus free-choice learning groups. We attribute at least some of this sample-wide benefit to the Transformative Experiences Model (Garner, Kaplan, & Pugh, 2016; Garner, Pugh, & Kaplan, 2016), as others have reported that even students participating in more structured learning experiences were able to glean personal relevance and meaning and expand their perception during this process (Ballantyne & Packer, 2006; Jackson, 1998). Rickinson (2001) suggested that learning programs in informal settings like museums have the potential to improve students' attitudes about learning. Additionally, other have found that transformative experiences can occur in the short-term (Garner, Kaplan, & Pugh, 2016; Koskey, Sondergeld, Stewart, & Pugh, 2018). College students enrolled in a course based on the Teaching for Transformative Experiences model reported being more interested and having higher academic performance than peers that did not participate (Heddy, Sinatra, Seli, Taasoobshirazi, & Mukhopadhyay, 2017). Pugh and Bergin (2005) proposed that the more intrinsically motivated and interested an individual is to learn or engage in some task, the more likely they are to undergo a transformative experience and potentially further develop their motivation and interest after this experience; thus, as all participants in our study indicated increased motivation and more positive feelings based on post-survey scores, it seems likely that at least some students were engaging in transformative experiences to a certain extent—whether at the zoo or in the classroom.

Limitations

We recognize that our advanced biology sample (n=17 students) was smaller than anticipated, and we had no complete data from an advanced free-choice learning group to complete our sampling design; thus, all participating advanced biology students were part of the structured learning experience. Future iterations of this research would benefit from comparison with a larger sample of advanced students and ones representing a free-choice advanced biology student group. Additionally, we recognize that we did not have a comparable non-zoo trip "control" group due to nearly all students across the three participating courses participating in the zoo trips; in the introductory biology and animal behavior courses, this trip was a required component of the class. However, as mentioned above, we did have one introductory student who was not able to participate in the scheduled zoo trip but did complete the pre- and post-surveys; more data could verify if this participant's trends mirrored students who might not attend a zoo trip. Again, while we cannot assume that the zoo trip wholly influenced all changes in motivation and positive feelings to learn biology among students over the semester, prior literature suggests this is very likely. Future iterations of this research would include control groups that would not participate in the zoo trip but would still complete the pre- and post-surveys.

Lastly, while we attempted to control for multiple factors in the structured learning visitor agendas, we could not guarantee an equal experience across all structured learning students due to unforeseen circumstances at the zoo (e.g., animal exhibits closed for cleaning or feeding, animal keeper demonstrations being cancelled or delayed, etc.), though our similar variances across treatment groups suggest this was not a concern.

Similarly, we could not guarantee an equal experience across all free-choice learning students due to the autonomous nature of the free-choice learning treatment.

Conclusions

All participating students—regardless of whether they were assigned to the structured or free-choice learning group, or were introductory or advanced biology students—reported changes in motivation and more positive feelings related to learning biology. Though we recognize these benefits may not fully be due to students' participation in the zoo trip—based on the absence of a control group—prior literature suggests benefits of learning experiences at informal settings. Ultimately, there may be numerous ways to make visits to the zoo—and presumably other informal settings like museums, aquariums, and science centers—more meaningful for undergraduates, whether instructors aim to offer more structured or autonomous learning experiences. However, future research including control groups will need to be conducted to confirm such trends.

CHAPTER VI

CONCLUSIONS REGARDING UNDERGRADUATE ENGAGEMENT IN BIOLOGY

The overarching goal of my dissertation research was to better understand how undergraduate students engage in biology. By studying undergraduates' learning experiences in the classroom (i.e., Part 1), I found that learner-centeredness in the college biology classroom is multidimensional, and often, that perceptions of those in the classroom as well as the metrics used to quantify learner-centeredness are misaligned. Specifically, the perceptions of student, instructor, and expert observers of learner-centeredness—based on an array of validated metrics—in a biology course were inconsistent. Thus, instructors should be aware of how their classroom practices are perceived by others, and how the various aspects of their courses could be made more learner-centered (Chapter II). Additionally, I found that the Decibel Analysis for Research in Teaching (DART; Owens et al., 2017) did not align well with validated learner-centered metrics such as the Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2002), and further, generally underestimated the learner-centeredness of a course session. As both instruments may be measuring discrete aspects of learner-centeredness, I suggest that additional research—including the inclusion of other learner-centered instruments and further validation of audio recording devices used with

DART—is necessary to wholly quantify learner-centered factors in the classroom (Chapter III).

By exploring undergraduates' learning experiences beyond the classroom at informal learning settings (i.e., Part 2), I discovered that informal learning experiences of biology undergraduates vary widely, and that such out-of-school experiences may be essential for both increasing student interest in biology and improving retention of students in undergraduate biology programs. I found that the Informal Learning Experiences Survey (ILES) may be most beneficial for practitioners in the classroom and program directors at informal learning settings as a means of better understanding the learning experiences of biology undergraduate students in an informal setting, rather than strictly as a research tool. Additionally, my survey results documented the number of barriers against participating in informal learning experiences and the number of reasons for learning about science among college-age adults related to the informal learning settings this age group regularly visits (Chapter IV).

I also concluded that all students demonstrated improvements in various types of motivation and positive feelings associated with learning biology based on pre- and post-survey scores, regardless of whether they were assigned to a structured or free-choice group, or whether they were introductory or advanced. Essentially, the level of structure incorporated into a learning experience at the zoo does not matter. Though we recognize these benefits may not fully be due to students' participation in the zoo trip—based on the absence of a control group—prior literature suggests benefits of learning experiences at informal settings (Chapter V).

Engaging undergraduates in biology through active methods of teaching and learning is essential for meaningful learning to occur (Fencl & Scheel, 2005). Instructors and program directors in biology must strive to alleviate the "unintentional loss" of students from science majors caused by more passive learning environments and instructional styles (Tanner & Allen, 2004). In light of the leaky STEM pipeline—in which reported attrition rates for students in science disciplines can approach nearly 50% (Chen & Soldner, 2013)—both reforming classrooms to be more learner-centered environments and including more learning experiences at informal settings have the potential to more fully engage undergraduate students in biology and improve retention rates of biology majors over time.

REFERENCES

Abdullah, M. Y., Bakar, N. R. A., & Mahbob, M. H. (2012). Student's participation in classroom: What motivates them to speak up? *Procedia-Social and Behavioral Sciences, 51*, 516-522.

Adams, J. D., & Branco, B. (2017). Extending classrooms into parks through informal science learning and place-based education. In Patrick P. (ed.), *Preparing Informal Science Educators* (337-354). New York, NY: Springer.

Adelman, L. M., Dierking, L. D., Haley Goldman, K., Coulson, D., Falk, J. H., & Adams, M. (2001). Baseline impact study: Disney's Animal Kingdom Conservation Station. *Technical Report*. Annapolis, MD: Institute for Learning Innovation.

Adelman, L. M., Falk, J. H., & James, S. (2000). Assessing the National Aquarium in Baltimore's impact on visitor's conservation knowledge, attitudes and behaviors. *Curator*, *43*(1), 33 – 62.

Allen, D., & Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. *Cell Biology Education, 4*(4), 262-268.

Alsop, S., & Watts, M. (1997). Sources from a Somerset village: A model for informal learning about radiation and radioactivity. *Science Education, 81*(6), 633-650.

Amrein-Beardsley, A., & Popp, S. E. O. (2012). Peer observations among faculty in a college of education: investigating the summative and formative uses of the Reformed Teaching Observation Protocol (RTOP). *Journal of Personnel Evaluation in Education, 24*(1): 5-24.

Andrews, T. M., Leonard, M. J., Colgrove, C. A., & Kalinowski, S. T. (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE—Life Sciences Education*, *10*(4), 394-405.

Antón, M. (1999). The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom. *The Modern Language Journal*, *83*(3), 303-318.

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE—Life Sciences Education*, *8*(3), 203-213.

Arya, D. J., & Maul, A. (2012). The role of the scientific discovery narrative in middle school science education: An experimental study. *Journal of Educational Psychology*, *104*(4), 1022.

Audacity Team (2017). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.2.1. Retrieved December 20, 2017, from: https://audacityteam.org/.

Ayeni, O. G., & Olowe, M. O. (2016). The implication of large class size in the teaching and learning of business education in tertiary institution in Ekiti State. *Journal of Education and Practice*, *7*(34), 65-69.

Ballantyne, R., Fien, J., & Packer, J. (2001). Program effectiveness in facilitating intergenerational influence in environmental education: Lessons from the field. *The Journal of Environmental Education*, *32*(4), 8-15.

Ballantyne, R., & Packer, J. (2006). Promoting learning for sustainability: Principals' perceptions of the role of outdoor and environmental education centres. *Australian Journal of Environmental Education*, *22*(1), 15-29.

Bandalos, D. L., & Finney, S. J. (2018). Factor analysis: Exploratory and confirmatory. In *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 110-134). Abingdon, UK: Routledge.

Barker, B. S., & Ansorge, J. (2007). Robotics as means to increase achievement scores in an informal learning environment. *Journal of Research on Technology in Education*, *39*(3), 229-243.

Barr, D.A., Gonzalez, M.E., and Wanat, S.F. (2008). The leaky pipeline: Factors associated with early decline in interest in premedical studies among underrepresented minority undergraduate students. *Academic Medicine, 83*(5): 503–511.

Baruch, Y. (1999). Response rate in academic studies-A comparative analysis. *Human Relations*, *52*(4), 421-438.

Beeth, M. E. (1998). Teaching for conceptual change: Using status as a metacognitive tool. *Science Education*, *82*(3), 343-356.

Bernstein, D. A. (2018). Does active learning work? A good question, but not the right one. *Scholarship of Teaching and Learning in Psychology*, *4*(4), 290.

Berryman, S., & McCune, B. (2006). Estimating epiphytic macrolichen biomass from topography, stand structure and lichen community data. *Journal of Vegetation Science*, *17*(2), 157-170.

Bevan, B., Dillon, J., Hein, G. E., Macdonald, M., Michalchik, V., Miller, D., Root, D., Rudder, L., Xanthoudaki, M., & Yoon, S. (2010). Making science matter: Collaborations between informal science education organizations and schools. *Washington, DC: Center for Advancement of Informal Science Education*.

Biggs, J., Kember, D., & Leung, D. Y. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, *71*(1), 133-149.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*(6), 740-756.

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3-4), 369-398.

Boekaerts, M., & Minnaert, A. (1999). Self-regulation with respect to informal learning. *International Journal of Educational Research*, *31*(6), 533-544.

Bogner, F. X. (1998). The influence of short-term outdoor ecology education on long-term variables of environmental perspective. *The Journal of Environmental Education*, *29*(4), 17-29.

Borg, C., & Mayo, P. (2005). The EU Memorandum on lifelong learning. Old wine in new bottles?. *Globalisation, Societies and Education*, *3*(2), 203-225.

Borun, M., Chambers, M. B., Dritsas, J., & Johnson, J. I. (1997). Enhancing family learning through exhibits. *Curator: The Museum Journal*, *40*(4), 279-295.

Bowl, M. (2001). Experiencing the barriers: Non-traditional students entering higher education. *Research Papers in Education*, *16*(2), 141-160.

Bradley, J. S. (2005). Does the classroom assist or impede the learning process? *Canadian Association of Principals Journal*, *13*, 32-34.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (eds.) and Committee on Developments in the Science of Learning, National Research Council (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academies Press.

Brown, P. L., Abell, S. K., Demir, A., & Schmidt, F. J. (2006). College science teachers' views of classroom inquiry. *Science Education*, *90*(5), 784-802.

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).

Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and… tensions with professional identity?. *CBE-Life Sciences Education*, *11*(4), 339-346.

Casagrand, J., & Semsar, K. (2017). Redesigning a course to help students achieve higher-order cognitive thinking skills: from goals and mechanics to student outcomes. *Advances in Physiology education*, *41*(2), 194-202.

Cattaneo, K. H. (2017). Telling active learning pedagogies apart: From theory to practice. *Journal of New Approaches in Educational Research (NAER Journal)*, *6*(2), 144-152.

Chen, X., & Soldner, M. (2013). *STEM Attrition: College Students' Paths into and out of STEM Fields*. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Chiou, G. L., Liang, J. C., & Tsai, C. C. (2012). Undergraduate students' conceptions of and approaches to learning in biology: A study of their structural models and gender differences. *International Journal of Science Education*, *34*(2), 167-195.

Chism, N. V. N. (2006). Challenging traditional assumptions and rethinking learning spaces. *Learning Spaces*, 2-1.

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, *45*(6), 378-387.

Cooper, K. M., & Brownell, S. E. (2016). Coming Out in Class: Challenges and Benefits of Active Learning in a Biology Classroom for LGBTQIA Students. *CBE-Life Sciences Education*, *15*(3), ar37.

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research, 77*(1), 113-143.

Covill, A. E. (2011). College students' perceptions of the traditional lecture method. *College Student Journal*, *45*(1), 92-102.

Crane, V., Nicholson, H., Chen, M., & Bitgood, S. (Eds.). (1994). *Informal Science Learning: What Research Says about Television, Science Museums, and Community-Based Projects*. Dedham, MA: Research Communication.

Crick, R., McCombs, B., Haddon, A., Broadfoot, P., & Tew, M. (2007). The ecology of learning: factors contributing to learner-centred classroom cultures. *Research Papers in Education*, *22*(3), 267-307.

Crosnoe, R., & Muller, C. (2014). Family socioeconomic status, peers, and the path to college. *Social Problems*, *61*(4), 602-624.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970-977.

Crowley, K., & Callanan, M. (1998). Describing and supporting collaborative scientific thinking in parent-child interactions. *Journal of Museum Education*, *23*(1), 12-17.

Csikszentmihalyi, M., & Hermanson, K. (1995). What makes visitors want to learn? Intrinsic motivation in museums. *Museum News*, *74*(3), 34-37.

Dabbagh, N., & Kitsantas, A. (2012). Personal learning environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education*, *15*(1), 3-8.

Dall'Alba, G., & Sandberg, J. (2006). Unveiling professional development: A critical review of stage models. *Review of Educational Research*, *76*(3), 383-412.

Daniels, D. H., Kalkman, D. L., & McCombs, B. L. (2001). Young children's perspectives on learning and teacher practices in different classroom contexts: Implications for motivation. *Early Education and Development, 12*, 253-273.

Dart, B. C., Burnett, P. C., Purdie, N., Boulton-Lewis, G., Campbell, J., & Smith, D. (2000). Students' conceptions of learning, the classroom environment, and approaches to learning. *The Journal of Educational Research*, *93*(4), 262-270.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, *332*(6031), 862-864.

Dierking, L. D. (2014). A view through another window: Free-choice science learning and generation R. In *Assessing Schools for Generation R (Responsibility)* (pp. 307-319). New York, NY: Springer.

Dierking, L. D., & Falk, J. H. (2009). Learning for life: The role of free-choice learning in science education. In *The World of Science Education* (pp. 179-205). Leiden, Netherlands: Brill Sense.

Donnelly, P. J., Blanchard, N., Olney, A. M., Kelly, S., Nystrand, M., & D'Mello, S. K. (2017, March). Words matter: automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 218-227). ACM.

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., Kelly, S., Nystran, M., & D'Mello, S. K. (2016, July). Automatic teacher modeling from live classroom audio. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (pp. 45-53). ACM.

Drissner, J. R., Haase, H. M., Wittig, S., & Hille, K. (2014). Short-term environmental education: long-term effectiveness? *Journal of Biological Education, 48*(1), 9-15.

Durham, M. F., Knight, J. K., Bremers, E. K., DeFreece, J. D., Paine, A. R., & Couch, B. A. (2018). Student, instructor, and observer agreement regarding frequencies of scientific teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO). *International Journal of STEM Education*, *5*(1), 31.

Eagan, K., Stolzenberg, E. B., Lozano, J. B., Aragon, M. C., Suchard, M. R., & Hurtado, S. (2014). *Undergraduate Teaching faculty: The 2013–2014 HERI Faculty survey*. Los Angeles: Higher Education Research Institute, UCLA.

Ebert-May, D., Derting, T. L., Henkel, T. P., Middlemis Maher, J., Momsen, J. L., Arnold, B., & Passmore, H. A. (2015). Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. *CBE—Life Sciences Education*, *14*(2), ar22.

Ebert-May, D., Derting, T.L., Hodder, J., Momsen, J.L., Long, T.M., & Jardeleza, S.E. (2011). What we say is not what we do: effective evaluation of faculty professional development programs. *BioScience, 61*(7), 550-558.

Eddy, S. L., Brownell, S. E., Thummaphan, P., Lan, M. C., & Wenderoth, M. P. (2015). Caution, student experience may vary: social identities impact a student’s experience in peer discussions. *CBE—Life Sciences Education*, *14*(4), ar45.

Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, *13*(3), 478-492.

Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: a classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE—Life Sciences Education*, *14*(2), ar23.

Edmunds, R., & Richardson, T. E. (2009). Conceptions of learning, approaches to studying and personal development in UK higher education. *British Journal of Educational Psychology, 79,* 295–309.

Ellenbogen, K. M., Luke, J. J., & Dierking, L. D. (2004). Family learning research in museums: An emerging disciplinary matrix? *Science Education*, *88*(S1), S48-S58.

Entwistle, N. (2008, April). Taking stock: teaching and learning research in higher education. In *Review Prepared for an International Symposium on "Teaching and Learning Research in Higher Education."* Guelph, Ontario, Canada.

Entwistle, N., McCune, V., & Hounsell, J. (2002). Approaches to studying and perceptions of university teaching-learning environments: Concepts, measures and preliminary findings. *Occasional Report*, *1*.

Ernst, H., & Colthorpe, K. (2007). The efficacy of interactive lecturing for students with diverse science backgrounds. *Advances in Physiology Education*, *31*(1), 41-44.

Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, *16*(2), 171-190.

Evans, C., Abrams, E., Reitsma, R., Roux, K., Salmonsen, L., & Marra, P. P. (2005). The Neighborhood Nestwatch Program: Participant outcomes of a citizen-science ecological research project. *Conservation Biology*, *19*(3), 589-594.

Fadigan, K. A., & Hammrich, P. L. (2004). A longitudinal study of the educational and career trajectories of female participants of an urban informal science education program. *Journal of Research in Science Teaching, 41*(8), 835-860.

Fahraeus, A. (2013). Research supports learner-centered teaching. *Journal of the Scholarship of Teaching and Learning*, 126-131.

Falconer, K., Joshua, M., Wyckoff, S., & Sawada, D. (2001, March). Effect of reformed courses in physics and physical science on student conceptual understanding. *Paper presented at the Annual Meeting of the National Association of Research in Science Teaching*. St. Louis, MO.

Falk, J., & Dierking, L. D. (2000). *Learning from Museums: Visitor Experiences and the Making of Meaning*. Walnut Creek, CA: AltaMira Press.

Falk, J., & Storksdieck, M. (2005). Using the contextual model of learning to understand visitor learning from a science center exhibition. *Science Education, 89*(5), 744-778.

Falk, J. H., & Adelman, L. (2003). Investigating the impact of prior knowledge, experience and interest on aquarium visitor learning. *Journal of Research in Science Teaching, 40*(2), 163-176.

Falk, J. H., & Dierking, L. D. (2002). *Lessons without Limit: How Free-Choice Learning is Transforming Education*. New York, NY: Rowman Altamira.

Falk, J. H., Dierking, L. D., & Foutz, S. (Eds.). (2007). *In Principle, in Practice: Museums as Learning Institutions*. Lanham, MD: Rowman Altamira.

Falk, J. H., & Gillespie, K. L. (2009). Investigating the role of emotion in science center visitor learning. *Visitor Studies, 12*(2), 112-132. doi: 10.1177/0963662506064240

Falk, J. H., & Needham, M. D. (2013). Factors contributing to adult knowledge of science and technology. *Journal of Research in Science Teaching, 50*(4), 431-452. doi:10.1002/tea.21080

Falk, J. H., Scott, C., Dierking, L., Rennie, L., & Jones, M. C. (2004). Interactives and visitor learning. *Curator, 47*, 171–192.

Falk, J. H., & Storksdieck, M. (2010). Science learning in a leisure setting. *Journal of Research in Science Teaching*, *47*(2), 194-212. doi:10.1002/tea.20319

Fassinger, P. A. (1996). Professors' and students' perceptions of why students participate in class. *Teaching Sociology*, 25-33.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Felder, R. M., and Brent, R. (1996). Navigating the bumpy road to student-centered instruction. *College Teaching 44*(2), 43-47.

Fencl, H., & Scheel, K. (2005). Engaging students. *Journal of College Science Teaching*, *35*(1), 20.

Floyd, K. S., Harrington, S. J., & Santiago, J. (2009). The effect of engagement and perceived course value on deep and surface learning strategies. *Informing Science: The International Journal of an Emerging Transdiscipline*, *12*, 181-190.

Fox-Cardamone, L., & Rue, S. (2003). Students' responses to active-learning strategies: An examination of small-group and whole-class discussion. *Research for Educational Reform, 8*(3), 3-15.

Fraser, B. J. (1994). Research on classroom and school climate. In D. Gabel (Ed.), *Handbook of Research on Science Teaching and Learning*. New York: MacMillan.

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

Gardner, H. (1991). *The Unschooled Mind*. New York, NY: Basic Books.

Garner, J. K., Kaplan, A., & Pugh, K. (2016). Museums as contexts for transformative experiences and identity development. *Journal of Museum Education, 41*(4), 341-352.

Garner, J. K., Pugh, K., & Kaplan, A. (2016). Museum visitor identification and engagement with science (VINES): A theory-driven process for designing transformational experiences. In *Annual Meeting of the American Educational Research Association, Washington DC*.

Gibbs, G., & Coffey, M. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education, 5*(1), 87-100.

Ginns, P., & Ellis, R. (2007). Quality in blended learning: Exploring the relationships between on-line and face-to-face teaching and learning. *The Internet and Higher Education*, *10*(1), 53-64.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176.

Gormally, C., Brickman, P., Hallar, B., & Armstrong, N. (2011). Lessons learned about implementing an inquiry-based curriculum in a college biology laboratory classroom. *Journal of College Science Teaching*, *40*(3).

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, *8*(3), 381-391.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology*, *92*(2), 316.

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, *100*(1), 105.

Harackiewicz, J. M., Tibbetts, Y., Canning, E., & Hyde, J. S. (2014). Harnessing values to promote motivation in education. *Advances in Motivation and Achievement: A Research Annual*, *18*, 71.

Harvey, E. J., & Kenyon, M. C. (2013). Classroom seating considerations for 21st century students and faculty. *Journal of Learning Spaces*, *2*(1).

Heddy, B. C., Sinatra, G. M., Seli, H., Taasoobshirazi, G., & Mukhopadhyay, A. (2017). Making learning meaningful: facilitating interest development and transfer in at-risk college students. *Educational Psychology*, *37*(5), 565-581.

Heim, A. B., & Holt, E. A. (2018). Comparing student, instructor, and expert perceptions of learner-centeredness in post-secondary biology classrooms. *PLoS ONE, 13*(7): e0200524. https://doi.org/10.1371/journal.pone.0200524

Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching*, *42*(5), 62-66.

Hilke, D. D., & Balling, J. D. (1985). The family as a learning system: An observational study of family behavior in an information rich environment. *Final Report Grant No.: SED-8112927, National Science Foundation, Washington, DC.*

Hofstein, A., & Rosenfeld., S. (1996). Bridging the gap between formal and informal science learning. *Studies in Science Education*, *28*, 87-112.

Holt, E. A., Young, C., Keetch, J., Larsen, S., & Mollner, B. (2015). The greatest learning return on your pedagogical investment: Alignment, assessment or in-class instruction?. *PLoS ONE, 10*(9), e0137446.

Howard, J. R., & Henney, A. L. (1998). Student participation and instructor gender in the mixed-age college classroom. *The Journal of Higher Education*, *69*(4), 384-405.

Huberman, M. (1981). *ECRI, Masepa, North Plains: A Case Study*. Andover, MA: The NETWORK, Inc.

HyperNiche. (2009). Nonparametric Multiplicative Habitat Modeling. Version 2.0. MjM Software, Gleneden Beach, Oregon, USA.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Jackson, P. W. (1998). *John Dewey and the Lessons of Art*. New Haven, CT: Yale University Press.

Kahl Jr., D. H., & Venette, S. (2010). To lecture or let go: A comparative analysis of student speech outlines from teacher-centered and learner-centered classrooms. *Communication Teacher*, *24*(3), 178-186.

Kantartzi, S. K., Allen, S., Lodhi, K., Grier IV, R. L., & Kassem, M. A. (2010). Study of factors affecting students' performance in three science classes: General biology, botany, and microbiology at Fayetteville State University. *Atlas Journal of Science Education*, *1*(1), 13-18.

Kaplan, A., Sinai, M., & Flum, H. (2014). Design-based interventions for promoting students' identity exploration within the school curriculum. *Motivational Interventions (Advances in Motivation and Achievement)*, *18*, 243-291.

Kassens-Noor, E. (2012). Twitter as a teaching practice to enhance active and informal learning in higher education: The case of sustainable tweets. *Active Learning in Higher Education, 13*(1), 9-21.

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis* (Vol. 344). Hoboken, NJ: John Wiley & Sons.

Kellert, S. R. (1997). *The Value of Life: Biological Diversity and Human Society*. Washington, DC: Island Press.

Kilday, C. R., & Kinzie, M. B. (2009). An analysis of instruments that measure the quality of mathematics teaching in early childhood. *Early Childhood Education Journal*, *36*(4), 365-372.

Kitsantas, A., & Dabbagh, N. (2011). The role of Web 2.0 technologies in self-regulated learning. *New Directions for Teaching and Learning, 2011*(126), 99-106.

Knight, J. K., Wise, S. B., & Southard, K. M. (2013). Understanding clicker discussions: student reasoning and the impact of instructional cues. *CBE—Life Sciences Education*, *12*(4), 645-654.

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

Kola-Olusanya, A. (2005). Free-choice environmental education: understanding where children learn outside of school. *Environmental Education Research*, *11*(3), 297-307.

Koran Jr., J. J., Koran, M. L., Foster, J. S., & Dierking, L. D. (1988). Using modeling to direct attention. *Curator: The Museum Journal*, *31*(1), 36-42.

Koskey, K. L., Sondergeld, T. A., Stewart, V. C., & Pugh, K. J. (2018). Applying the mixed methods instrument development and construct validation process: The transformative experience questionnaire. *Journal of Mixed Methods Research*, *12*(1), 95-122.

Kranzfelder, P., Bankers-Fulbright, J. L., García-Ojeda, M. E., Melloy, M., Mohammed, S., & Warfa, A. R. M. (2019). The Classroom Discourse Observation Protocol (CDOP): A quantitative method for characterizing teacher discourse moves in undergraduate STEM learning environments. *PLoS ONE*, *14*(7), e0219019.

Lasry, N., Charles, E., Whittaker, C., Dedic, H., & Rosenfield, S. (2013, January). Changing classroom designs: Easy; Changing instructors' pedagogies: Not so easy... In *AIP Conference Proceedings* (Vol. 1513, No. 1, pp. 238-241). American Institute of Physics.

Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, *44*(5), 706-724.

Lee, M. H., Johanson, R. E., & Tsai, C. C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science Education*, *92*(2), 191-220.

Lei, S. A. (2010). Classroom physical design influencing student learning and evaluations of college instructors: A review of literature. *Education*, *131*(1), 128-135.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, *97*(2), 184.

Li, Y., & Dorai, C. (2006). Instructional video content analysis using audio information. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(6), 2264-2274.

Loh Epri, M. (2016). A case study on the impact of large classes on student learning. *Contemporary PNG Studies*, *24*, 95.

Lombardi, M. M., & Wall, T. B. (2006). *Learning Spaces*, 17-1. Duke University: Perkins Library.

MacIsaac, D., Sawada, D., & Falconer, K. (2001). Using the Reformed Teaching Observation Protocol (RTOP) as a catalyst for self-reflective change in secondary science teaching. Paper presented at *Annual Meeting of the American Educational Research Association, Seattle, Washington*.

MacIsaac, D., & Falconer, K. (2002). Reforming physics instruction via RTOP. *The Physics Teacher*, *40*(8), 479-485.

Madge, C., Meek, J., Wellens, J., & Hooley, T. (2009). Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work'. *Learning, Media and Technology*, *34*(2), 141-155.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954.

Marshall, J. C., Smart, J., Lotter, C., & Sirbu, C. (2011). Comparative analysis of two inquiry observational protocols: Striving to better understand the quality of teacher-facilitated inquiry-based instruction. *School Science and Mathematics*, *111*(6), 306-315.

Martell, S. T. (2008). Of cultural tools and kinds of knowledge about art, culture, and the environment: Investigating field trip-based learning. *Journal of Museum Education*, *33*(2), 209-220.

Martin, B. L., Mintzes, J. J., & Clavijo, I. E. (2000). Restructuring knowledge in biology: cognitive processes and metacognitive reflections. *International Journal of Science Education*, *22*(3), 303-323.

Marton, F., Hounsell D., & Entwistle, N. (Eds.) (1997) *The Experience of Learning*. Edinburgh: Scottish Academic Press.

Marton, F. M., & Säljö, R. (2005). Approaches to learning. In F. M. Marton, D. Hounsell, & N. J. Entwistle (Eds.), *The Experience of Learning: Implications for Teaching and Studying in Higher Education* (3rd ed., pp. 106–125). Edinburgh: Centre for Teaching, Learning and Assessment, University of Edinburgh.

Mayo, M. J. (2009). Video games: A route to large-scale STEM education?. *Science*, *323*(5910), 79-82.

McCombs, B. (2003, April). Defining tools for teacher reflection: The assessment of learner-centered practices. Paper presented at the *Annual Meeting of the American Educational Research Association*.

McCombs, B. L. (2000). Assessing the role of educational technology in the teaching and learning process: A learner-centered perspective. Paper presented at *The Secretary's Conference on Educational Technology, 2000: Shaping the Future*, *Alexandria, VA.*

McCombs, B. L., & Miller, L. (2007). *Learner-Centered Classroom Practices and Assessments: Maximizing Student Motivation, Learning, and Achievement*. Thousand Oaks, CA: Corwin Press.

McCombs, B. L., & Quiat, M. (2002). What makes a comprehensive school reform model learner centered?. *Urban Education*, *37*(4), 476-496.

McCune, B., & Mefford, M. J. (2016). PC-ORD. Multivariate Analysis of Ecological Data. Version 7. MjM Software Design, Gleneden Beach, Oregon, U.S.A.

McKeachie, W., & Svinicki, M. (2013). *McKeachie's Teaching Tips*. Boston, MA: Cengage Learning.

Meesuk, K., & Srisawasdi, N. (2014). Implementation of student-associated game-based open inquiry in chemistry education: Results on students' perception and motivation. In *Proceedings of the 22nd International Conference on Computers in Education* (pp. 219-226).

Miller, C. J., & Metz, M. J. (2014). A comparison of professional-level faculty and student perceptions of active learning: its current use, effectiveness, and barriers. *Advances in Physiology Education*, *38*(3), 246-252.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, *85*(3), 424.

Modell, H. I., & Michael, J. A. (1993). *Promoting Active Learning in the Life Science Classroom*. New York Academy of Sciences.

Montgomery, T. (2008). Space matters: Experiences of managing static formal learning spaces. *Active Learning in Higher Education*, *9*(2), 122-138.

Murray, K., & MacDonald, R. (1997). The disjunction between lecturers' conceptions of teaching and their claimed educational practice. *Higher Education*, *33*(3), 331-349.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews, 82*(4), 591-605.

Naqvi, A., Venugopal, B., Falk, J. H., & Dierking, L. D. (1991). Family visitors to the National Museum of Natural History, New Delhi: A research analysis. *Curator, 34*(1), 51–57.

National Research Council (NRC). (1996). *Learning Science in Informal Environment*s. Washington, DC: National Academies Press.

National Research Council (NRC). (2009). *Learning Science in Informal Environment*s. Washington, DC: National Academies Press.

Negra, C., & Manning, R. E. (1997). Incorporating environmental behavior, ethics, and values into nonformal environmental education programs. *The Journal of Environmental Education*, *28*(2), 10-21.

Oblinger, D. G. (2006). Space as a change agent. *Learning Spaces*, *1*.

Olson, J. K., Cox-Petersen, A. M., & McComas, W. F. (2001). The inclusion of informal environments in science teacher preparation. *Journal of Science Teacher Education*, *12*(3), 155-173.

O'Neill, G., & Guerin, S. (2015). Working with the challenge of designing and implementing a stand-alone learning to learn module in a large Arts programme. *AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education, 7*(3).

Owens, M. T., Seidel, S. B., Wong, M., Bejines, T. E., Lietz, S., Perez, J. R., ... & Balukjian, B. (2017). Classroom sound can be used to classify teaching practices in college science courses. *Proceedings of the National Academy of Sciences*, 201618693.

Owens, M. T., Trujillo, G., Seidel, S. B., Harrison, C. D., Farrar, K. M., Benton, H. P., ... & Byrd, D. T. (2018). Collectively improving our teaching: attempting biology department–wide professional development in scientific teaching. *CBE—Life Sciences Education*, *17*(1), ar2.

Paris, S. G. (1997). Situated motivation and informal learning. *Journal of Museum Education*, *22*(2-3), 22-27.

Park, E. L., & Choi, B. K. (2014). Transformation of classroom spaces: Traditional versus active learning classroom in colleges. *Higher Education*, *68*(5), 749-771.

Pearson, J. C., & West, R. (1991). An initial investigation of the effects of gender on student questions in the classroom: Developing a descriptive base. *Communication Education*, *40*(1), 22-32.

Piburn, M., & Sawada, D. (2000). Reformed Teaching Observation Protocol (RTOP) reference manual. *Technical Report*.

Pike, G. R., & Kuh, G. D. (2005). First-and second-generation college students: A comparison of their engagement and intellectual development. *The Journal of Higher Education*, *76*(3), 276-300.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*(4), 667.

Pintrich, P. R., & DeGroot, E. (1990, April). Quantitative and qualitative perspectives on student motivational beliefs and self-regulated learning. In *Annual Meeting of the American Educational Research Association, Boston, MA* (Vol. 128).

Pintrich, P. R., & Garcia, T. (2012). Self-regulated learning in college students: Knowledge, strategies, and motivation. In *Student Motivation, Cognition, and Learning* (pp. 129-150). London, UK: Routledge.

Pratt, D. D., & Collins, J. B. (2000). The Teaching Perspectives Inventory (TPI). Paper presented at the *Adult Education Research Conference*, Vancouver, BC.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of English Education 93*, 223-231.

Pugh, K. J. (2004). Newton's laws beyond the classroom walls. *Science Education*, *88*(2), 182-196.

Pugh, K. J., & Bergin, D. A. (2005). The effect of schooling on students' out-of-school experience. *Educational Researcher*, *34*(9), 15-23.

Pugh, K. J., Linnenbrink-Garcia, L., Koskey, K. L., Stewart, V. C., & Manzey, C. (2010). Motivation, learning, and transformative experience: A study of deep engagement in science. *Science Education*, *94*(1), 1-28.

Randol, S. (2004). Looking for inquiry: Developing an instrument to assess inquiry at museum exhibits. In *Annual meeting of the American Educational Research Association (AERA). San Diego, California*.

Rennie, L. J., & Johnston, D. J. (2004). The nature of learning and its implications for research on learning from museums. *Science Education*, *88*(S1), S4-S16.

Richardson, J. C., & Newby, T. (2006). The role of students' cognitive engagement in online learning. *The American Journal of Distance Education*, *20*(1), 23-37.

Rickinson, M. (2001). Learners and learning in environmental education: A critical review of the evidence. *Environmental Education Research*, *7*(3), 207-320.

Rideout, B. E. (2005). The effect of a brief environmental problems module on endorsement of the new ecological paradigm in college students. *The Journal of Environmental Education*, *37*(1), 3-11.

Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education*, *59*(2), 185-213.

Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, *81*(4), 493-529.

Rushton, G. T., Lotter, C., & Singer, J. (2011). Chemistry teachers' emerging expertise in inquiry teaching: The effect of a professional development model on beliefs and practice. *Journal of Science Teacher Education*, *22*(1), 23-52.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*(5), 749.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sachatello-Sawyer, B., & Fellenz, R. (2000). Coming of age: A national study of adult museum programs. *Curator: The Museum Journal*, *43*(2), 147-156.

Sachatello-Sawyer, B., Fellenz, R., Burton, H., Gittings-Carlson, L., Lewis-Mahony, J., & Woolbaugh, W. (2002). *Adult Museum Programs: Designing Meaningful Experiences*. Lanham, MD: Rowman Altamira.

Sambell, K., Brown, S., & McDowell, L. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, *23*(4), 349-71.

Sanders, M. (2013). Classroom design and student engagement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2013 (57),* 496.

Sansone, C., & Harackiewicz, J. M. (Eds.). (2000). *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*. Cambridge, MA: Academic Press.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). Reformed Teaching Observation Protocol (RTOP). *ACEPT Technical Report No. IN00-1*. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics, 102*(6), 245-253.

Schuetze, H. G., & Slowey, M. (2002). Participation and exclusion: A comparative analysis of non-traditional students and lifelong learners in higher education. *Higher Education*, *44*(3-4), 309-327.

Schwan, S., Grajal, A., & Lewalter, D. (2014). Understanding and engagement in places of science experience: Science museums, science centers, zoos, and aquariums. *Educational Psychologist, 49*(2), 70-85.

Seep, B., Glosemeyer, R., Hulce, E., Linn, M., & Aytar, P. (2000). *Classroom Acoustics: A Resource for Creating Environments with Desirable Listening Conditions*. Melville, NY: Acoustical Society of America Publications.

Seymour, E., & Hewitt, E. (1997). *Talking About Leaving: Factors Contributing to High Attrition Rates Among Science, Mathematics, and Engineering Undergraduate Majors*. Boulder, CO: Bureau of Sociological Research.

Shavelson, R. J., Webb, N. M. & Burstein, L. (1986). Measurement of teaching. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching,* (pp. 50–91). New York, NY: Macmillan.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Singer, J., Lotter, C., Feller, R., & Gates, H. (2011). Exploring a model of situated professional development: Impact on classroom practice. *Journal of Science Teacher Education*, *22*(3), 203-227.

Skogsberg, K., & Clump, M. (2003). Do psychology and biology majors differ in their study processes and learning styles?. *College Student Journal, 37*(1), 27.

Smith, M. K., Jones, F. H., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. *CBE-Life Sciences Education*, *12*(4), 618-627.

Stuckey, B., & Arkell, R. (2006). Development of an e-learning knowledge sharing model. A report for the *Knowledge Sharing Services Project, Australian Flexible Learning Framework Supporting E-Learning Opportunities*. Accessed from: http://pre2009.flexiblelearning.net.au/flx/go/home/projects/2005/pid/48.

Su, H., Dzodzo, B., Wu, X., Liu, X., & Meng, H. (2019). Unsupervised methods for audio classification from lecture discussion recordings. *Proceedings of Interspeech 2019*, 3347-3351.

Subramaniam, M. A. (2002). *Garden-Based Learning in Basic Education: A Historical Review*. Center for Youth Development, University of California. Retrieved from: http://ucce.ucdavis.edu/freeform/4hcyd/documents/ CYD Monograph (Focus)2609.pdf.

Sutton, R. E., & Wheatley, K. F. (2003). Teachers' emotions and teaching: A review of the literature and directions for future research. *Educational Psychology Review*, *15*(4), 327-358.

Tabachnick, B. R., & Zeichner, K. M. (1999). Idea and action: Action research and the development of conceptual change teaching of science. *Science Education*, *83*(3), 309-322.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273-1296.

Tanner, K., & Allen, D. (2004). Approaches to biology teaching and learning: learning styles and the problem of instructional selection—engaging all students in science courses. *Cell Biology Education*, *3*(4), 197-201.

Tanner, K. D. (2013). Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education*, *12*(3), 322-331.

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education*, 247-280.

Theobald, R., & Freeman, S. (2014). Is it the intervention or the students? Using linear regression to control for student characteristics in undergraduate STEM education research. *CBE-Life Sciences Education*, *13*(1), 41-48.

Tiwari, A., Lam, D., Yuen, K. H., Chan, R., Fung, T., & Chan, S. (2005). Student learning in clinical nursing education: Perceptions of the relationship between assessment and learning. *Nurse Education Today*, *25*(4), 299-308.

Trigwell, K. (2002). Approaches to teaching design subjects: a quantitative analysis. *Art, Design & Communication in Higher Education*, *1*(2), 69-80.

Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review, 16*(4), 409-424.

Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education*, *37*(1), 57-70.

Tsai, C. C. (2004). Conceptions of learning science among high school students in Taiwan: A phenomenographic analysis. *International Journal of Science Education*, *26*(14), 1733-1750.

Tsang, A., & Harris, D. M. (2016). Faculty and second-year medical student perceptions of active learning in an integrated curriculum. *Advances in Physiology Education*, *40*(4), 446-453.

Tudor, J., Penlington, R., & McDowell, L. (2010). Perceptions and their influences on approaches to learning. *Engineering Education*, *5*(2), 69-79.

van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing*, *25*(4), 40.

Veltri, S., Banning, J. H., & Davies, T. G. (2006). The community college classroom environment: Student perceptions. *College Student Journal*, *40*(3), 517-528.

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: integrating active learning into a large lecture course. *CBE-Life Sciences Education*, *7*(4), 361-367.

Wandersee, J. H., Mintzes, J. J., & Novak, J. D. (1994). Research on alternative conceptions in science. *Handbook of Research on Science Teaching and Learning*, *177*, 210.

Wang, Z., Pan, X., Miller, K. F., & Cortina, K. S. (2014). Automatic classification of activities in classroom discourse. *Computers & Education*, *78*, 115-123.

Watters, D. J., & Watters, J. J. (2007). Approaches to learning by students in the biological sciences: Implications for teaching. *International Journal of Science Education*, *29*(1), 19-43.

Weimer, M. (2002). *Learner-Centered Teaching: Five Key Changes to Practice*. Hoboken, NJ: John Wiley & Sons.

Weinberger, E., & McCombs, B. L. (2003). Applying the LCPs to high school education. *Theory into Practice, 42*(2), 117-126.

Wellington, J. (1990). Formal and informal learning in science: The role of the interactive science centres. *Physics Education*, *25*(5), 247.

Wendel, P., Young, A., Esson, J., & Plank, K. (2016, March). Motivational decline in higher education STEM courses and its association with student demographics. In *Conference Proceeding. New Perspectives in Scienze Education* (pp. 203-207).

Wentzel, K. R., & Brophy, J. E. (2014). *Motivating Students to Learn*. Abingdon, UK: Routledge.

White, B.Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Science*, *16*, 90-91.

Wieman, C., & Gilbert, S. (2014). The teaching practices inventory: A new tool for characterizing college and university teaching in mathematics and science. *CBE-Life Sciences Education*, *13*(3), 552-569.

Wigfield, A., Klauda, S. L., & Cambria, J. (2011). Influences on the development of academic self-regulatory processes. *Handbook of Self-Regulation of Learning and Performance*, 33-48.

Wilbur, T. G., & Roscigno, V. J. (2016). First-generation disadvantage and college enrollment/completion. *Socius*, *2*, 2378023116664351.

Wilde, M. (2007). Das contextual model of learning—ein theorierahmen zur erfassung von lernprozessen in museen. In *Theorien in der Biologiedidaktischen Forschung* (pp. 165-175). Berlin, Germany: Springer.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: a test of self-determination theory. *Journal of Personality and Social Psychology*, *70*(4), 767.

Wolins, I. S., Jensen, N., & Ulzheimer, R. (1992). Children's memories of museum field trips: A qualitative study. *Journal of Museum Education*, *17*(2), 17-27.

Woodin, T., Carter, V. C., & Fletcher, L. (2010). Vision and change in biology undergraduate education, a call for action—initial responses. *CBE—Life Sciences Education*, *9*(2), 71-73.

Yamamura, S., & Takehira, R. (2017). Effect of practical training on the learning motivation profile of Japanese pharmacy students using structural equation modeling. *Journal of Educational Evaluation for Health Professions*, *14*.

Zagallo, P., Meddleton, S., & Bolger, M. S. (2016). Teaching real data interpretation with models (TRIM): Analysis of student dialogue in a large-enrollment cell and developmental biology course. *CBE—Life Sciences Education*, *15*(2), ar17.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64-70.

Zimmerman, H. T., & McClain, L. R. (2015). Family learning outdoors: Guided participation on a nature walk. *Journal of Research in Science Teaching*, *53*(6), 919-942.

APPENDIX A

INSTITUTIONAL REVIEW BOARD
APPROVAL LETTERS

Appendix A1: IRB Approval Letter for Chapters II and III.

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

DATE:	November 6, 2017
TO:	Emily Holt
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[932641-2] Teaching and Learning in BIOL 1010
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	November 6, 2017
EXPIRATION DATE:	May 31, 2020

Thank you for your submission of Amendment/Modification materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

Appendix A2: IRB Approval Letter for Chapter IV.

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

DATE:	April 23, 2018
TO:	Ashley Heim
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[1227292-2] Improving the Informal Learning Experiences Survey (ILES)
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	April 23, 2018
EXPIRATION DATE:	April 23, 2022

Thank you for your submission of Amendment/Modification materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

Appendix A3: IRB Approval Letter for Chapter V.

UNIVERSITY OF
NORTHERN COLORADO

Institutional Review Board

DATE:	August 21, 2018
TO:	Ashley Heim
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[1301825-1] Understanding Undergraduates' Informal Learning Experiences at the Denver Zoo
SUBMISSION TYPE:	New Project
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	August 20, 2018
EXPIRATION DATE:	August 19, 2022

Thank you for your submission of New Project materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Ashley -

Thank you for a very thorough and clear IRB application for an interesting research project.

All materials and protocols are verified/approved exempt. Two minor items needs to be addressed in your consent forms before they are used in your protocols:

1) Please replace Sherry May's name with Nicole Morse in all consent forms as Sherry retired at the end of June 2018 and is no longer the IRB administrator or contact person for participant mistreatment reporting/contact.

2) Please also be sure to emphasize that the decision whether or not to participate in this research will have NO negative or positive consequences to evaluation in BIO 110 in the consent forms.

Best wishes with this research and don't hesitate to contact me with any IRB-related questions or concerns.

Sincerely,

Dr. Megan Stellino, UNC IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Nicole Morse at 970-351-1910 or nicole.morse@unco.edu. Please include your project title and reference number in all correspondence with this committee.

APPENDIX B

INFORMAL LEARNING EXPERIENCES
SURVEY VERSION 1

Appendix B

<u>Item 1.</u> How many times have you visited the following places or engaged in the following activities in the **last 6 months**? (Indicate # of visits for each option using the sliding bar. Bars range from 0-10+.)

1. Zoo	
2. Aquarium	
3. Museum	
4. Nature/Environmental Center	
5. Science Center	
6. Space Center/Planetarium	
7. Nature Preserve/Conservancy	
8. Botanical Gardens	
9. State/National Park	
10. Local Nature Areas/Trails/City Parks	
11. Educational Club (on- or off-campus)	
12. Educational Camp	
13. Other	

If you indicated "Other" above, please describe your response.

<u>Item 2.</u> Why do you visit the above locations to learn about science? (Choose all that apply.)

1. To learn about something new.
2. To gather with friends and family.
3. To explore a new area/location.
4. To further understand/review topics learned in the classroom.
5. Nostalgia, I visited these areas as a child or teenager.
6. Special events occurring at these places (e.g. traveling exhibits at a museum, seasonal exhibits at the zoo, concerts at botanical gardens).
7. It was required as part of a class or work.
8. These experiences are designed for my age group.
9. I feel culturally and socially accepted at these places.
10. I volunteer at one or more of these places.
11. Other (If you indicated “Other,” please describe your response.)

Appendix B, Continued.

Item 3. What barriers prevent you from visiting the above places as often as you would prefer? (Choose all that apply.)

1. Limited finances/experiences are too expensive.
2. Lack of transportation/experiences are too far away.
3. School responsibilities.
4. Job responsibilities.
5. Family responsibilities.
6. Not interested or motivated to participate.
7. Not aware of special events occurring at these places (e.g. traveling exhibits at a museum, seasonal exhibits at the zoo).
8. Experiences not designed for my age group.
9. I don't feel culturally or socially accepted at informal learning institutions.
10. Other (If you indicated "Other," please describe your response.)

Please choose the most accurate response for each of the following statements.
When visiting the above locations to learn about science…

Not at all true (1)	(2)	(3)	Somewhat true (4)	(5)	(6)	Very true (7)

Item 4. I enjoy participating in these science learning experiences very much.
Item 5. These science learning experiences are fun.
Item 6. I think these science learning experiences are boring.
Item 7. Science learning experiences do not hold my attention at all.
Item 8. I would describe science learning experiences as very interesting.
Item 9. I think science learning experiences are quite enjoyable.
Item 10. While participating in science learning experiences, I think about how much I enjoy them.

Item 11. With whom do you generally participate in informal learning experiences? (Choose all that apply.)

1. Siblings
2. Parents
3. Children
4. Significant other
5. Other family members (If you indicated "Other family member," please describe your response.)
6. Friends from school
7. Friends from outside of school
8. I prefer to go by myself
9. Other (If you indicated "Other," please describe your response.)

Item 12. Which of the following places did you visit as a child or teenager? (Choose all that apply.)

1. Zoo
2. Aquarium
3. Museum
4. Nature/Environmental Center
5. Science Center
6. Space Center/Planetarium
7. Nature Preserve/Conservancy
8. Botanical Gardens
9. State/National Parks
10. Local Nature Areas/Trails/City Parks
11. Other (If you indicated "Other," please describe your response.)

APPENDIX C

DESCRIPTIVE PARAMETERS OF FACTORS FOR ITEMS ON INFORMAL LEARNING EXPERIENCES SURVEY VERSION 1 BASED ON EXPLORATORY FACTOR ANALYSIS AND ITEM RELIABILITY

Appendix C

Item	Description	Factors	Mean	Standard Deviation	Reliability Estimate
1	Frequency/type of informal learning	1	1.30	1.967	0.919
		2	1.53	1.609	0.782
2	Reasons for learning about science	1	0.30	0.426	0.329
		2	0.07	0.251	0.356
3	Barriers	1	0.44	0.452	0.397
		2	0.53	0.481	0.288
4	People	1	0.46	0.457	0.485
		2	0.23	0.394	0.163
		3	0.57	0.475	0.302
5	Informal learning as children/teens	1	0.71	0.405	0.758

APPENDIX D

INFORMAL LEARNING EXPERIENCES
SURVEY VERSION 2

Appendix D

<u>Item 1.</u> How many times have you visited the following places or engaged in the following activities in the **last 6 months**?

	0	1	2	3	4	5	6	7	8	9	10 or more
1. Zoo or Animal Sanctuary											
2. Aquarium											
3. Museum											
4. Science Center or Butterfly Pavilion (focuses specifically on science, usually indoors)											
5. Nature Center/Preserve (generally focuses on biology or other outdoor sciences, usually outdoors)											
6. Space Center/Planetarium (focuses on astronomy or other space-related sciences, usually indoors)											
7. Botanical Gardens											
8. City/State/National Parks (including nature areas and trails in these locations)											
9. Theme Parks with a specific focus on science/conservation (examples: Disney's Animal Kingdom, Sea World, etc.)											
10.Educational Club (includes clubs you've been involved at school, in your community, for your church, etc., with a focus on science education)											
11. Educational Camps (includes camps you've attended for school, in your community, for your church, etc., with a focus on science education)											
12. Other											

If you indicated "Other" above, please describe your response.

Appendix D, Continued.

Item 2. Why do you currently visit the above locations to learn about science? (Choose all that apply.)

12. To learn about something new.
13. To gather with friends and family.
14. Just for fun. I find the experience enjoyable.
15. To explore a new area/location.
16. To further understand/review topics learned in the classroom.
17. Nostalgia, I visited these areas as a child or teenager.
18. Special events occurring at these places (e.g. traveling exhibits at a museum, seasonal exhibits at the zoo, concerts at botanical gardens).
19. It was required as part of a class or work.
20. These experiences are designed for my age group.
21. I feel culturally and socially accepted at these places.
22. I volunteer at one or more of these places.
23. Other (If you indicated "Other," please describe your response.)

Item 3. What barriers prevent you from visiting the above places as often as you would prefer? (Choose all that apply.)

11. Limited finances/experiences are too expensive.
12. Lack of transportation/experiences are too far away.
13. School responsibilities, including extracurricular school activities.
14. Job responsibilities.
15. Family responsibilities.
16. Social responsibilities (e.g. hanging out with friends).
17. Not interested or motivated to participate.
18. Not aware of special events occurring at these places (e.g. traveling exhibits at a museum, seasonal exhibits at the zoo).
19. Experiences not designed for my age group.
20. I don't feel culturally or socially accepted at informal learning institutions.
21. Other (If you indicated "Other," please describe your response.)

Item 4. With whom do you generally participate in informal learning experiences? (Choose all that apply.)

10. Siblings
11. Parents
12. Friends
13. Children
14. Significant other
15. Teachers/mentors
16. I prefer to go by myself
17. Other (If you indicated "Other," please describe your response.)
18. Other family members (If you indicated "Other family members," please describe your response.)

Item 5. Which of the following places did you visit as a child or teenager? (Choose all that apply.)

12. Zoo or Animal Sanctuary
13. Aquarium
14. Museum
15. Science Center or Butterfly Pavilion (focuses specifically on science, usually indoors)
16. Nature Center/Preserve (generally focuses on biology or other outdoor sciences, usually outdoors)
17. Space Center/Planetarium (focuses on astronomy or other space-related sciences, usually indoors)
18. Botanical Gardens
19. City/State/National Parks (including nature areas and trails in these locations)
20. Theme Parks with a specific focus on science/conservation (examples: Disney's Animal Kingdom, Sea World, etc.)

Appendix D, Continued.

21. Educational Club (includes clubs you've been involved at school, in your community, for your church, etc., with an educational focus)
22. Educational Camps (includes camps you've attended for school, in your community, for your church, etc., for a specific educational purpose)
23. Other (If you indicated "Other," please describe your response.)

APPENDIX E

INFORMAL LEARNING EXPERIENCES SURVEY DEMOGRAPHIC QUESTIONNAIRE

Appendix E

You will now be asked to respond to several demographic questions in the last portion of this survey. Please answer all questions honestly and to the best of your ability.

Item 1. What is your race/ethnicity? Select all that apply.

- ☐ American Indian or Native Alaskan
- ☐ Asian
- ☐ Black or African American
- ☐ Hispanic or Latino/a/x
- ☐ Native Hawaiian or other Pacific Islander
- ☐ White
- ☐ Unknown
- ☐ My race/ethnicity is not listed.
- ☐ Prefer not to state.

Item 2. Which best describes your gender identity?

- ○ Woman
- ○ Man
- ○ Transgender Woman
- ○ Transgender Man
- ○ Gender-Queer or Gender-Nonconforming
- ○ Questioning
- ○ My identity is not listed.
- ○ Prefer not to state.

Item 3. Indicate the course for which you took this survey for extra credit.

- ○ Introductory principles of biology
- ○ Introductory organismal biology

Item 4. What is your estimated grade in the course selected above?

- ○ A
- ○ B
- ○ C
- ○ D
- ○ F

Appendix E, Continued.

Item 5. What is your major? (Drop-down menu format)

- Biological Science: Pre-health and Biomedical Science emphasis
- Biological Science: Ecology and Evolution
- Biological Science: Cell and Molecular Biology emphasis
- Biological Science: Secondary Teaching
- Audiology and Speech-Language Sciences
- Chemistry
- Dietetics
- Earth Sciences
- Environmental and Sustainability Studies
- Mathematics
- Nursing
- Physics
- Psychology
- Sport and Exercise Science
- Undecided
- Other

If you indicated "Other" for your major above, please type your major here.

Item 6. In what year of study are you at the institution where this research was conducted? (Drop-down menu format)

- 1st year
- 2nd year
- 3rd year
- 4th year
- 5th year
- 6th year or beyond

Item 7. Are you a transfer student?

- Yes
- No

Appendix E, Continued.

Item 8. What is your age?

- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26 or older

Item 9. What is your best guess as for the yearly income of the household in which you grew up?

- $0 to $50,000
- $50,000-$100,000
- More than $100,000
- Decline to state
- Do not know

Item 10. What is your mother's highest level of education?

- Did not finish high school
- GED
- High School Diploma
- Technical Degree/Certificate
- Associate's Degree
- Bachelor's Degree
- Master's Degree
- Doctoral Degree
- Unknown
- Other

If you indicated "Other" above, please describe your response here.

Appendix E, Continued.

Item 11. What is your father's highest level of education?

- Did not finish high school
- GED
- High School Diploma
- Technical Degree/Certificate
- Associate's Degree
- Bachelor's Degree
- Master's Degree
- Doctoral Degree
- Unknown
- Other

If you indicated "Other" above, please describe your response here.

Item 12. What is your marital status?

- Single, never married
- In a relationship, never married
- Married or domestic partnership
- Widowed
- Divorced
- Separated

Item 13. How many children do you have?

- 0
- 1
- 2
- 3 or more

Item 14.

If you are from the U.S., what is the zip code of the town/city where you spent the most time growing up?

If you from outside the U.S., what are the city and country where you spent the most time growing up?

APPENDIX F

INTRODUCTORY STRUCTURED ZOO CONTENT ASSESSMENT

Appendix F. While three versions of this handout were developed which aligned with the three "structured learning group" visitor agendas described in the methods, all versions include the same items, just in a different order. Therefore, only one version is included here. Additionally, exhibit names have been altered slightly to maintain anonymity of the zoo visited in this study.

1.**Giraffe exhibit**: Name one physical or physiological adaptation the giraffe has developed for living in a savanna habitat.

2. **Fish/amphibians/reptile exhibit**: Which of the following animal clades are present in this exhibit? Circle all that apply.

Cyclostomata	Chondrichthyes	Actinopterygii	Amphibia
Testudines	Lepidosauria	Crocodilia	Aves
Mammalia			

3. **Fish/amphibians/reptile exhibit**: For each characteristic listed below, provide the name of an organism you observed in this exhibit that possesses that characteristic. Common names rather than species names are okay. (Try to come up with different organisms for each characteristic!)

Bilateral symmetry: ______________________

Undergoes ecdysis (skin-shedding): ______________________

Closed circulatory system: ________________________

Gills for respiration and excretion: ________________________

Reproduces via external fertilization: ________________________

Deuterostome development: ___________________________

Bony skeleton: __________________________

Four paired-limbs (tetrapod): __________________________

Epidermal scales: _________________________

4. **Big cats exhibit:** After observing the large cats on display, describe 3 adaptations that tigers (or other large predatory cats) have developed for a carnivorous diet.

5. **Bird exhibit:** As you walk through this exhibit, name at least 2 adaptations that birds have developed for flight.

What is the African penguin's main mode of locomotion? ___________________

Name 2 adaptations that the penguin has for this form of locomotion.

Appendix F, Continued.

6. **Primate exhibit**: Are primates considered amniotes? YES / NO

Explain your response.

7. **Primate exhibit**: What is one diagnostic characteristic of primates and other members of Class Mammalia? (Note: Diagnostic means the feature is found in all members of an animal clade and ONLY in that one animal clade.)

8. **Zookeeper Talks:**

Choose one of the zookeeper talks you attended today: ______________________________

What is one thing you learned about animal behavior specific to the species discussed in the zookeeper talk?

APPENDIX G

FREE-CHOICE ZOO CONTENT ASSESSMENTS

Appendix G1. Introductory Free-choice Zoo Content Assessment. Exhibit names have been altered slightly to maintain anonymity of the zoo visited in this study.

On the following map, circle the exhibits that you visit during your trip. Please include arrows to show the order in which you visit these exhibits. Next to each exhibit you circle, please provide an estimate of how long you spent at the exhibit.

[map removed to maintain anonymity of zoo]

Please respond to the following questions (in any order you choose) during your zoo trip today. As you respond to these questions, here are some key terms and concepts from organismal biology to keep in mind:

Terms
Actinopterygii
Amniota
Amphibia
Aves
Bilateral symmetry
Bony skeleton
Chondrichthyes
Circulatory systems (open vs. closed)
Crocodilia
Cyclostomata
Deuterostome
Ecdysis (skin-shedding)
Epidermal scales
Fertilization (external vs. internal)
Gills for respiration and excretion
Lepidosauria
Mammalia
Protostome
Testudines
Tetrapoda

Appendix G1, Continued.

Concepts

Diagnostic characteristic means the feature is found in all members of an animal clade and ONLY in that one animal clade.

Consider how different species are adapted for their particular habitats.

Consider how different species are adapted for their particular diets.

---

1. What was the most interesting piece of information you learned today relating to animal biology?

2. Considering what you know about organismal diversity, do you think this zoo has a diverse enough selection of animals on exhibit? Explain your response.

What clade(s) would you recommend this zoo include in their exhibits to increase diversity of species at the zoo? Explain your response.

3. What was your favorite animal/exhibit that you visited today? Explain why.

4. What was your favorite thing that you experienced today that was NOT an animal/exhibit (e.g., zoo atmosphere, food, rides, shops, etc.)?

5. Did you attend of the zookeeper talks (or any other special demonstration) during your visit today? YES / NO

If you responded YES, please answer the following questions:

List the talks/demonstrations that you attended.

What is one thing you learned about animal behavior specific to the species discussed in the zookeeper talk/demonstration?

Why did you choose to attend these talks/demonstrations?

Appendix G2. Advanced Free-choice Zoo Content Assessment. Exhibit names have been altered slightly to maintain anonymity of the zoo visited in this study.

On the following map, circle the exhibits that you visit during your trip. Please include arrows to show the order in which you visit these exhibits. Next to each exhibit you circle, please provide an estimate of how long you spent at the exhibit.

[map removed to maintain anonymity of zoo]

Please respond to the following questions (in any order you choose) during your zoo trip today.

1. What was the most interesting piece of information you learned today relating to mammalogy?

2. Considering what you know about mammal taxa, do you think this zoo has a diverse enough selection of mammals on exhibit? Explain your response.

What clade(s) would you recommend this zoo include in their exhibits to increase diversity of mammal species at the zoo? Explain your response.

3. What was your favorite animal/exhibit that you visited today? Explain why.

4. What was your favorite thing that you experienced today that was NOT an animal/exhibit (e.g., zoo atmosphere, food, rides, shops, etc.)?

5. Did you attend of the zookeeper talks (or any other special demonstration) during your visit today? YES / NO

If you responded YES, please answer the following questions:

List the talks/demonstrations that you attended.

What is one thing you learned about animal behavior specific to the species discussed in the zookeeper talk/demonstration?

Why did you choose to attend these talks/demonstrations?

APPENDIX H

PRIOR EXPERIENCES AT ZOOS QUESTIONNAIRE

Appendix H

1. When was the last time you visited a zoo?
   A. Within the last month
   B. Within the last 6 months
   C. Within the last year
   D. 1-2 years ago
   E. 2-3 years ago
   F. Greater than 3 years ago
   G. I have never visited a zoo.

2. Specifically, when was the last time you visited this zoo?
   A. Within the last month
   B. Within the last 6 months
   C. Within the last year
   D. 1-2 years ago
   E. 2-3 years ago
   F. Greater than 3 years ago
   G. I have never visited this zoo.

3. How often do you visit zoos each year (on average)?
   A. 1 time
   B. 2 times
   C. 3 times
   D. 4 times
   E. 5 or more times
   F. I generally never visit zoos.

4. How often do you visit this zoo each year (on average)?
   A. 1 time
   B. 2 times
   C. 3 times
   D. 4 times
   E. 5 or more times
   F. I generally never visit zoos.

APPENDIX I

LEARNING SELF-REGULATION QUESTIONNAIRE

Appendix I

The following questions relate to your reasons for participating actively on your zoo trip. Different people have different reasons for their participation in informal learning experiences, and we want to know how true each of the reasons is for you. Please use the following scale to indicate how true each reason is for you:

1 (Not at all true) 2 3 4 (Somewhat true) 5 6 7 (Very true)

1. I will participate actively on the zoo trip:
   A. Because I feel like it's a good way to improve my understanding of biology material.
   B. Because others might think badly of me if I didn't.
   C. Because I would feel proud of myself if I learned something on the trip.
   D. Because a solid understanding of biology is important to my intellectual growth.

2. I am likely to follow my instructor's suggestions for what to do on the zoo trip:
   A. Because I would get a bad grade if I didn't do what he/she suggests.
   B. Because I am worried that I am not going to perform well in the course.
   C. Because it's easier to follow his/her suggestions than come up with my own learning strategies.
   D. Because he/she seems to have insight about how best to learn biology material.

3. The reason that I will work to expand my knowledge of biology on the zoo trip is:
   A. Because it's interesting to learn more about the nature of biology.
   B. Because it's a challenge to really understand how to solve problems in biology.
   C. Because a good grade in biology will look positive on my record.
   D. Because I want others to see that I am intelligent.

APPENDIX J

PERSONAL INTEREST IN BIOLOGY METRIC

Appendix J

Instructions: For each question, select the response that best matches the extent to which you agree or disagree.

[Responses were on a 5-point Likert-scale, Strongly Disagree to Strongly Agree]

Value

1. Biology concepts are valuable because they will help me in the future.
2. Biology concepts are practical for me to know.
3. Biology concepts will be useful for me later in life.
4. Biology concepts help me in my daily life outside of school.

Feeling

1. I enjoy Biology.
2. I am fascinated by Biology.
3. I like Biology.
4. The field of Biology is exciting to me.