School of Education

A Flipped Classroom Approach in Undergraduate Anatomy and Physiology: A Mixed Methods Study Evaluating Learning Environment and Student Outcomes

Scott Edward Gaines

0000-0001-6234-4691

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

July 2021

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number #HRE2018-0468, and the Concordia University Institutional Review Board (IRB).

Signature: ………………………………………….

Date: ………………………...

Abstract

Undergraduate human anatomy and physiology (A&P) is considered a cornerstone course in biomedical education. However, there is currently a concern that A&P classes are not properly preparing students for future clinical coursework. One pedagogical tool that shows promise as a potential solution by changing classroom dynamics is the flipped classroom approach. While the topic of flipped classrooms has been prominent in educational literature in the past two decades, most studies have focused strictly on learning outcomes/academic achievement of students. Flipped classrooms, however, create an entirely new learning paradigm for the student, emphasizing active learning in the classroom. The purpose of this explanatory sequential mixed-methods study is to explore how the implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students.

Data was collected in two different classes of a second semester anatomy and physiology course at a small liberal arts university. One class (the treatment) utilized a flipped classroom structure while the other (the control) utilized a traditional lecture format. Quantitative data in the form of surveys were collected first. The College and University Classroom Environment Inventory (CUCEI) was administered to assess the students' perceptions of their learning environment. The Test of Science-Related Attitudes (TOSRA) was administered to assess students' attitudes toward science. Learning outcomes data was also compared in the form of lecture and laboratory practical examination grades. Qualitative data were then collected in the form of semi-structured focus groups and instructor reflections. After initial analysis and triangulation of quantitative and qualitative data was conducted, more quantitative data were collected in the form of student GPA data.

Quantitative data were analyzed using the Mann-Whitney U test for unpaired data and the Wilcoxon signed-rank test for paired data. Effect sizes were calculated using non-parametric tests to estimate the percentage of variance in the dependent variable that is explained by the independent variable. There were significantly higher scores in the flipped classroom on three scales of the CUCEI: innovation ($p = 0.04$), involvement ($p = 0.01$), and individualization ($p = 0.03$), with moderate to large effect sizes. There were no significant differences found in either the TOSRA or the outcomes data. Qualitative data was analyzed and coded in Quirkos™ qualitative analysis software. Qualitative data confirmed that the majority students in the flipped classroom enjoyed the format. It also became clear that certain vocal students in both classrooms were resistant to the concept. Instructor reflections specifically recalled high performing students being resistant to the idea of the flipped classroom. While there was no significant difference between those who liked the flipped classroom model and those that did not, the top students did fall into the latter category.

This study complements previous flipped classroom research and demonstrates that the flipped classroom does change the learning environment of undergraduate classrooms. Qualitative data brought to light the fact that some students have negative preconceptions regarding the flipped model. In order to be successful, instructors of the flipped method need to clearly articulate the rationale and benefits for the model in order to get student acceptance and participation.

Acknowledgements

This study has been made possible by the contribution and support of many people. My sincerest and heartfelt gratitude goes out to everyone who has supported me on this journey. The past several years has been incredibly challenging, but through the assistance, advice and encouragement of these individuals, this thesis was able to come to fruition.

I especially want to thank my supervisors, Professor John Williams and Professor Barry Fraser for their patience and their guidance. Without their help, this academic milestone would not be possible. So many times, I was unsure of myself and their calming words of insightful feedback would be able to bring me back on track. I'm sure my continual emails and requests for zoom meetings must have been trying for them, but they always responded graciously with words of sage advice. For everything you've done, I'm greatly indebted to you both.

I would also like to thank my coworkers at Concordia University Irvine. Specifically, my coworkers in the Biology department including Dr. Sean Bignami, Dr. Lindsay Kane-Barnese, Dr. Sarah Karam, Professor Marcella Olave, Dr. Roderick Soper, Dr. Mary Von dem Bussch, and Professor Tyler Zarubin. Each of them has given me sound advice and encouragement in this process. Special thanks to Drs. Soper and Von dem Bussch who assisted in the data collection process. In addition, thanks to Dr. Soper and fellow Concordia professor, Dr. Bret Taylor, who introduced me to the field of Learning Environment research and to Curtin University.

I also want to thank Dr. Guy White and his book, "The Dissertation Warrior: The Ultimate Guide to Being the Kind of Person Who Finishes a Doctoral Dissertation or Thesis." His insights were invaluable.

None of this would have been capable if it weren't for the loving support of my family. My mother, Patricia Gaines, who along with being a great cheerleader, also helped keep me organized while I worked two jobs and collected data. I also want to thank my wonderful children, James, John, Samantha, and Sara, who had to put up with me while I completed this project, including times I had to lock myself away in my office.

Most significantly, I want to thank my best friend and wife of 23 years. Kelli, you've been there through all my frustrations and late nights. You took on extra load of parenthood so I could have time to work, all the while working as a teacher yourself. There is no way I could have completed this without you. I can never thank you enough for all you do for me. Please know that your support, more than any other, has allowed me to reach my goal. Thank you.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Background to the Study

I have been an anatomy and physiology instructor at a small, Lutheran, liberal arts university for the past 8 years. I have seen first-hand how students struggle with the material, both in volume and complexity. Prior to teaching at the postsecondary level, for many years I worked (and continue to consult) for a fitness education company. There I oversee the development of health, fitness, nutrition and performance curriculum for various fitness professionals (e.g., personal fitness trainers, group exercise instructors, athletic coaches, etc.). Over my many years with the fitness education company, the demand shifted from students wanting to attend live training to on-demand e-learning. Although I initially doubted the efficacy of e-learning, I became well-versed in the various forms of e-learning program development and soon became a believer that, done well, online instruction could achieve the learning outcomes just as well as face-to-face training in certain situations (at times, even better).

Prior to ever hearing the term, “flipped classroom” I would often use my e-learning skills to develop assignments/lectures for my undergraduate students to do in their own time. I did this as a way to assist in covering the large volume of material and to open up class time to discussion. Anecdotally, some students seem to really enjoy the new online component and others would not. Likewise, some students who typically struggled seem to do better when I included the online components. Once I was introduced to the concept of the flipped classroom and saw how it was similar to what I had already done in small doses (just more structured), I became interested to see if this method could overcome the common challenges I saw in my classroom.

Later, a colleague introduced me to the field of learning environments. As I researched flipped classrooms and undergraduate anatomy and physiology instruction, no research seemed to be able to give the complete picture. Classroom environment surveys seemed to be the missing piece (to add to traditional assessment of learning outcomes) to evaluate the dynamics of a flipped undergraduate anatomy and physiology classroom.

1.2 Challenges to the Undergraduate Anatomy and Physiology Classroom

It has been stated that undergraduate human anatomy and physiology courses are not properly preparing students for careers in the medical sciences (Eseonu, Carachi, & Brindley, 2013; Johnston et al., 2015; Older, 2004; Raftery, 2007). In addition, while anatomy and physiology serves as foundation studies for multiple degrees in the health sciences, they typically have some of the highest failure rates of all undergraduate health science courses (Vitali, Blackmore, Mortazavi, & Anderton, 2020). As one of the most content-dense and conceptually challenging courses in the health sciences curriculum, it is reported that even those who do pass often barely meet minimum knowledge levels for their industry (Johnston et al., 2015). This often continues through post-graduate work, where students struggle due to only having a cursory introduction to anatomy as an undergraduate, and report a lack of confidence in their anatomical knowledge (Priyadharshini, Kumar, & Devi, 2019; Raftery, 2007). These difficulties and high failure rates in undergraduate anatomy and physiology can delay student entry into post-graduate work and various career pathways, create lower morale (of both students and instructors), increase costs (for the students and the institution), and often leave students with a negative perception of science (Anderton, Chiu, et al., 2016; Gultice, Witham, & Kallmeyer, 2015). In order to best understand the nature of the problem, one must first examine the nature of an undergraduate anatomy and physiology course and its place in biomedical education.

1.2.1 Undergraduate Anatomy and Physiology Courses

For many, undergraduate human anatomy and physiology courses are considered to be the cornerstone of biomedical curriculum (Priyadharshini et al., 2019). Their importance to the medical sciences has been recognized throughout history as noted scientists and philosophers such as Hippocrates, Aristotle, Galen, Avicenna, da Vinci, and Vesalius observed the importance of understanding human structure and function to treat disease (Eseonu et al., 2013; Saladin, 2018). In modern clinical fields this knowledge is recognized as pivotal for competent practice (El-Sayed & El-Sayed, 2013; Priyadharshini et al., 2019) as well as an indicator for future academic performance (S. J. Brown, White, & Power, 2017). As such, undergraduate human anatomy and physiology courses are a keystone subject for advanced coursework in the clinical sciences (Cliff & Wright, 1996). They are the common prerequisite course for virtually all medical related degrees including, but not limited to: pre-med, nursing, pre-physical therapy, pre-physician assistant, pre-dental, radiography, speech and language therapy, exercise science, biochemistry and biology (Stetzik, Deeter, Parker, & Yukech, 2015; Vitali et al., 2020). The goal of the anatomy and physiology instructor should be not only to teach content but also engage the learner and enhance the critical thinking skills needed in a clinical setting (El-Sayed & El-Sayed, 2013).

It is common for students to struggle in these courses due to the unique nature of the human anatomy and physiology content. Obviously, there are many variables that contribute to student success, and every student is unique, bringing their own distinct backgrounds, perspectives, and learning styles. However, there are some commonalities within human anatomy and physiology classrooms that make it especially difficult for instructors to guide students toward success. Most prevalent, anatomy and physiology instructors are faced with the problem of how to help students to learn and retain a very large volume of complex information and recall the pertinent facts in a clinical setting (Cliff & Wright, 1996; Dobson, 2013). More specifically, as explained by Vitali et. al. (2020), “the study of human anatomy requires significant intellectual effort to identify the diverse structures and their internal organisation, as well as comprehension of their relationships with other structures of the body” (p. 289). This often leads students to be intimidated by what is often seen as one of the most challenging courses within their curriculum. Furthermore, if students do not have a strong science background and have had a negative previous experience in the biosciences, they can become “science-phobic.”(Johnston et al., 2015, p. 416) Therefore, for students such as this, the challenging areas of human anatomy and physiology need to be presented in a manner that helps students overcome anxiety and support repetition/memorization learning and the conceptual understanding required (Johnston et al., 2015).

The need for students to be able to critically apply a large volume of information in human anatomy and physiology courses creates a paradox for many due to improper learning strategies. Due to the large amount of content, many students have the misconception that they must focus on memorization over reasoning skills. Put another way, students often misconstrue memorization for understanding (Miller, Perrotti, Silverthorn, Dalley, & Rarey, 2002). However, it has been argued that undergraduate students need to have a paradigm shift away from surface learning approaches that emphasize memorization, toward deeper learning strategies "characterized by a drive to understand underlying principles and concepts by grappling meaningfully with content" (McLean, Attardi, Faden, & Goldszmidt, 2016, p. 47). Furthermore, the diverse background of students within a typical undergraduate anatomy and physiology class (e.g., major, socioeconomic, preparedness for university, etc.) often leads to students with different challenges that the instructor must be able to address (Anderton, Chiu, et al., 2016; Higgins-Opitz & Tufts, 2015).

1.2.2 Undergraduate Anatomy and Physiology Instruction

However, undergraduate instructors, while being subject matter experts in their field, are typically not educated in the science of instruction (Mattheis & Jensen, 2014). Ideally, proper instructional design should lead to increased instructional efficiency and improved student learning, however, most anatomy and physiology instructors, or STEM (Science, Technology, Engineering and Math) instructors as a whole, have received little-to-no training in the science of instructional design and tend to teach as they themselves were taught (in a didactic/lecture-based format) (Khalil & Elkhider, 2016; Mattheis & Jensen, 2014). It has been argued that this teacher-centric method places the conveyance of information over problem solving as well as the development of critical thinking skills that is necessary for clinical practice (Mattheis & Jensen, 2014; Stetzik et al., 2015).

Instructors must also be able to adapt to an ever-expanding biomedical science curriculum. This, in turn, decreases the allotted time to study anatomy and physiology, giving instructors in the curriculum more to teach while the students often become overwhelmed. In fact, it has been said that the time dedicated to gross anatomical instruction has decreased more than 50% compared to 40 years ago (Eseonu et al., 2013; Johnston et al., 2015; Older, 2004; Vitali et al., 2020). This shortened time allotted for the study of anatomy and physiology also can have a secondary consequence of students failing to appreciate the relevance of the subject, further encouraging superficial learning strategies (Vitali et al., 2020).

1.2.3 The Flipped Classroom Model

A potential solution for struggling anatomy and physiology students is for instructors to implement a flipped classroom model. In a flipped (also referred to hybrid or inverted) classroom approach, content is introduced to the students prior to a scheduled class, thereby allowing students to engage with the content at a deeper level within the classroom, often in a more applied manner (Strayer, 2012). Put another way, what is normally done as homework is done in class and what is normally done in class is done as homework, hence the term "flipped" (Herreid, C. F., & Schiller, 2013). However, in practice, the flipped classroom is more than a simple re-ordering of classroom and homework activities. Rather, classroom activities are intended to be student-centric (as opposed to a typical teacher-centric lecture model). Active learning modalities are the main focus within the classroom, with group-based activities being an important component of the process (Bishop & Verleger, 2013). There is not a singular model for a flipped classroom, hence making a strict definition elusive. To create an all-encompassing working definition for their work on the flipped classroom approach, Abeysekera and Dawson (2015) proposed that a flipped classroom be defined as "a set of pedagogical approaches that:

> 1) move most information-transmission teaching out of class;
>
> 2) use class time for learning activities that are active and social and;
>
> 3) require students to complete pre- and/or post-class activities to fully benefit from in-class work" (Abeysekera & Dawson, 2015, p. 3).

The information presented prior to class may be presented in multiple forms, such as notes with images, augmented reality, and video lectures (the most common form) (Ferrer-Torregrosa et al., 2016). Hence, this "approach repurposes class time to focus on application and discussion; the acquisition of basic concepts and principles is done on the students' own time before class" (McLean et al., 2016, p. 47). This would seem to be an ideal solution for an anatomy and physiology classroom because many anatomy educators agree that success depends on overcoming the student misconception that they must focus on memorization and, instead, promoting a focus on reasoning skills to understand anatomy (Miller et al., 2002).

The general problem is that, despite the growing popularity of the flipped classroom model, researchers have yet to fully demonstrate the overall utility of the flipped classroom model in an anatomy and physiology classroom. This may partially be due to the fact that it is a relatively new instructional model, with some of the earliest papers discussing "inverting the classroom" being published only 20 years ago (Lage, Platt, & Treglia, 2000). However, research in the flipped classroom model has been very active since its inception (in August 2020, the term "flipped classroom" yielded a ProQuest search of nearly 12,000 results and Google Scholar search found almost 50,000). Hence, there seems to be great interest in the method and its potential benefits to demand such growth in research in such a relatively short period of time. However, its utility is still in question in part, due to the lack of strict definition or structure to the flipped classroom approach leading to ambiguity in its implementation, as well as a lack of any consensus on the learning theories which should serve as its foundation (Bishop & Verleger, 2013; J. L. Jensen, Holt, Sowards, Heath Ogden, & West, 2018). This ambiguity, in addition to the unique characteristics of the anatomy and physiology classroom and the fact that the flipped classroom creates a whole new learning experience for the student, inspired me to take a holistic view to study the effects of implementing the method; one that takes into account the three major domains of learning.

To fully evaluate any instructional approach such as the flipped classroom method, there are three domains of learning to be examined: the cognitive domain, the psychomotor domain and the affective domain. The cognitive domain relates to the area of learning concerned with intellectual ability such as knowledge and thinking skills (Aubrey & Riley, 2019). The focus of the cognitive domain is learning outcomes and is often the sole domain evaluated in educational studies. The psychomotor domain relates to the acquisition of practical or physical skills and strategy learning. (Aubrey & Riley, 2019; Yang & Tsai, 2012). An example of this in an anatomy and physiology class is the ability to make a proper incision when dissecting a specimen. The affective domain relates to the area of learning relating to feelings, emotions, behaviors, and beliefs (Aubrey & Riley, 2019; Yang & Tsai, 2012). Unlike the cognitive and psychomotor domains which both focus on learning, the focus of the affective domain is perceptions and attitudes.

1.2.4 The Study of Learning Environments

Traditional measures of learning outcomes do not paint a complete picture of the educational process (Fraser, 2003). Much of the research on instructional strategies in an anatomy and physiology classroom focus solely on learning outcomes which emphasize the cognitive and psychomotor domains. Likewise, often research on flipped classrooms does not consider the affective domain (Day, 2018; O'Flaherty & Phillips, 2015). As a relatively new educational tool, and one that is still developing, students may not have been exposed to the flipped classroom model. Hence, implementing the flipped classroom completely reorganizes the learning process, creating a new learning paradigm for the undergraduate anatomy and physiology student, restructuring the classroom environment. It has been shown that students' perceptions of their classroom environment is a potent determinant of student outcomes (Fraser, 2012), therefore evaluating their perceptions (i.e., the affective domain) should not be ignored if one's goal is to improve the effectiveness of a class.

The field of learning environments research is a well-established field that studies students' (and sometimes teachers') perceptions of psychosocial characteristics of their classroom which has been shown to have strong association with both cognitive and affective learning outcomes. Fraser (2012) defines classroom environment "in terms of the shared perceptions of the students and sometimes the teachers in that environment" (p. 1). The roots of classroom environment research are typically attributed to the work of Walberg and Moos in the early 1970s (Fraser, 2012). Since its early beginnings nearly 50 years ago, science educators worldwide have created and validated a large range of robust instruments to assess classroom environments at all levels (primary schools to postsecondary). The contributions of these researchers have created greater awareness that improved student outcomes can be achieved by creating positive classroom environments (Fraser, 2003). While some study has been conducted examining the changes in the classroom learning environment with the implementation of a flipped classroom (Strayer, 2012), to date none have addressed implementing the flipped approach within the unique structure and requirements of an undergraduate anatomy and physiology classroom.

1.3 Problem Statement

The specific problem is that researchers have not examined how the flipped classroom affects all three cognitive domains in a human anatomy and physiology classroom. Researchers have established that some flipped classrooms can positively affect some undergraduate science classrooms (Bergmann & Sams, 2012; Day, 2018; Khanova, Roth, Rodgers, & McLaughlin, 2015). It is unknown to researchers how the flipped classroom will affect an undergraduate anatomy and physiology classroom due to its unique dynamics. Therefore, researchers are calling for more research regarding the effect of implementing the flipped classroom method on student perceptions of their learning environment and their attitudes towards science. Additionally, researchers also are looking for new strategies to assist struggling anatomy and physiology students, with the flipped classroom being looked upon as a potential solution.

1.4 Purpose Statement

The problem statement above came from my background and early research into the fields of undergraduate anatomy and physiology education, flipped classrooms and learning environments. From the problem statement, the purpose statement below was drafted.

The purpose of this explanatory sequential mixed-methods study is to explore how the implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students.

1.5 Research Questions

As described in the purpose statement above, this study intends to examine if the changes in classroom dynamics, inherent in the flipped classroom method, are beneficial to undergraduate human anatomy and physiology students by examining their perceptions of the learning environment, their attitudes towards the subject, and their learning outcomes. As such, the following primary research question was created to articulate the goal of the project with secondary research questions generated to create measurable constructs.

Primary research question: *What is the effect of a flipped classroom environment on undergraduate students in a human anatomy and physiology course?*

Secondary research questions: *How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students'*:

1. *perceptions of their learning environment?*
2. *attitudes toward science?*
3. *examination achievement?*
4. *practical laboratory experience?*

The primary research question will be addressed by integrating qualitative and quantitative data sources in an explanatory sequential mixed methods format. Quantitative data via examinations, lab practicals, quizzes and GPA were collected to assess student outcomes (addressing the latter two of the secondary research questions in the cognitive and psychomotor domains). Further quantitative data were collected to assess the former two secondary questions (attitudes and learning environment perceptions, i.e., components of the affective domain) by well-established surveys. Specifically, the CUCEI (College and University Classroom Environment Inventory) (Fraser, Treagust, & Dennis, 1986) and TOSRA (Test of Science-Related Attitudes) (Fraser, 1978) were utilized. Qualitative data in the form of focus groups and instructor reflections helped to bring clarity and insight to the quantitative data.

1.6 Conceptual Framework

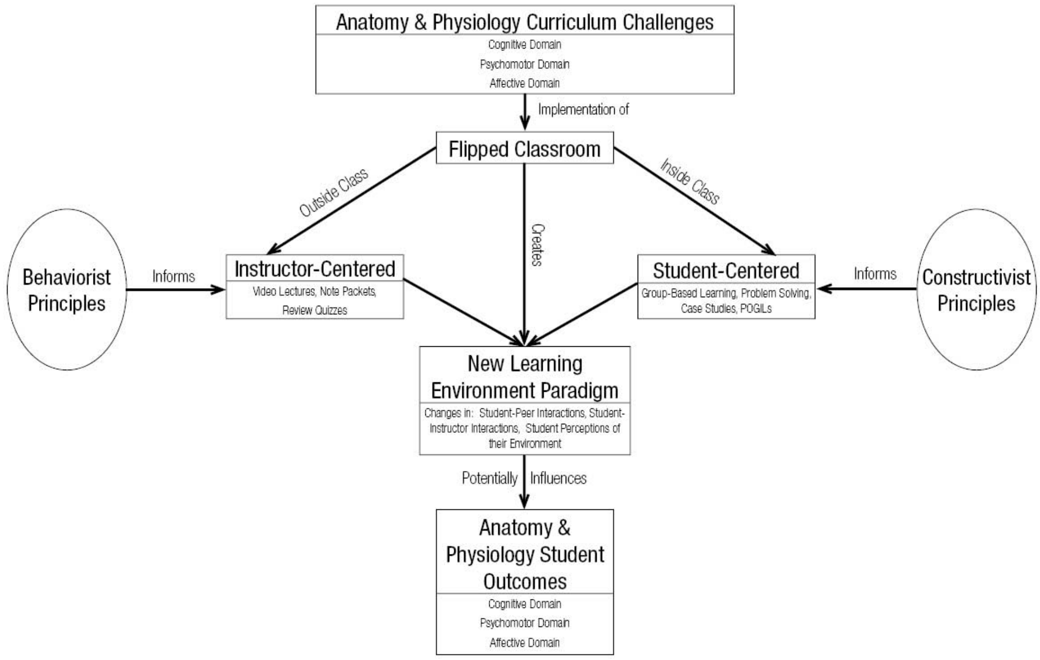


Figure 1-1 Conceptual Framework of the study

Figure 1 is a visual representation of the conceptual framework for this study. It shows how the implementation of the flipped classroom, utilizing both student-centered (inside class) and instructor-centered (outside class) modalities are intended to address some of the typical challenges of an anatomy and physiology classroom in all three learning domains (cognitive, psychomotor, and affective). This, in turn, creates a whole new learning environment for the students and instructor alike. The rationale for this framework will be detailed in Section 2.2 and its connection to the methodology in Section 3.2.2.

1.7 Significance of the Study

Determining the impact of a flipped classroom approach on an undergraduate science course such as human anatomy and physiology has potential to change education in the medical/clinical sciences. The significance of this study can be seen at multiple levels. First and foremost, is the significance to the study of flipped classrooms in science education. By analyzing not only outcomes but also student attitudes, I hope to get a clearer picture of which variables contribute to the success of the flipped classroom model in an undergraduate science classroom and which do not.

Second is the significance to instructors. The flipped-classroom approach changes the role of the instructor from lecturer to guide, with much greater student-teacher interaction. Ideally, the flipped classroom approach should allow teachers to focus on individual student weaknesses, misconceptions and the practical aspects of the study of human anatomy and physiology which are essential for careers in the medical sciences. In addition, this model should also show an increase in peer-to-peer interactions. Qualitative interviews will allow the researcher to find out if students perceive these increases in inter-personal interactions to be a benefit.

Next is the potential significance to science department administrators. Resources can be dedicated to the production and continual upkeep of flipped anatomy and physiology courses should student outcomes and attitudes be found to warrant it, or redirected to other areas if not warranted. Furthermore, the model utilized has the potential to be used in other science courses, especially those with direct clinical applications.

Lastly is the potential significance to the field of learning environments. To date, there have been relatively few learning environment studies on flipped classrooms and this will further the field of learning environments in this popular instructional strategy.

1.8 Overview of the Thesis

The first chapter in this thesis described the general problems leading to the rationale, purpose, conceptual framework and significance of my study.

The second chapter consists of a detailed literature review of the relevant topics in this study. It begins with a detailed explanation of the conceptual framework of the study. Next, a historical review of the common problems facing undergraduate anatomy and physiology classrooms is discussed. It then details the development of the flipped classroom approach to instruction and its potential strengths and weaknesses. This is followed by a review of the study of learning environments, its history and development of common surveys, and a detailed review of the two surveys utilized in this study - the CUCEI (College and University Classroom Environment Inventory) and TOSRA (Test of Science-Related Attitudes).

Chapter 3 describes the research methodology utilized in this study. The chapter begins with an overview of the conceptual components and philosophical foundation for the study. This leads to an explanation of the research design and specifically for the decision to use a mixed methods approach to address the research questions. The sample is described as well as the development of the curriculum utilized based on the conceptual framework of the study. Data collection methods for both quantitative and qualitative data are described as well as the statistical analysis techniques used to analyze them. Next, the methods to integrate both forms of data described in order to give a clearer picture of the subjects' experience. Lastly, potential limitations to the study are acknowledged and the methods used to minimize threats to reliability and validity are described.

The fourth chapter reports the data analysis and results of the study. Quantitative data are first analyzed, followed by qualitative data and the integration of both forms of data through triangulation. The chapter then discusses the results and how they address the research questions.

The fifth and final chapter summarizes the findings, discusses the study's findings and makes suggestions for future research.

Chapter 2. Literature Review

2.1 Introduction

The purpose of this explanatory sequential mixed-methods study is to explore how implementation of the flipped classroom format impacted the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students. The genesis for my study was based on my personal experience as an undergraduate anatomy and physiology instructor as well as an e-learning instructional designer. To my knowledge, it is the first study of its kind to integrate the flipped classroom technique in an undergraduate anatomy and physiology classroom and assess student perceptions of this learning environment in addition to traditional outcomes assessment.

This chapter will provide a detailed review of the literature regarding the core topics of the study. It begins with a description of the conceptual framework of the study (Section 2.2), reviewing each of the key components. Section 2.3 reviews the history of anatomy and physiology education, its importance to clinical settings (Section 2.3.1), and ultimately leading to an explanation of the most common struggles students and instructors have (Section 2.3.2). Section 2.4 will introduce the flipped classroom method. The historical background of this relatively new instructional method (Section 2.4.1), with beginnings dating back only 20 years, will be reviewed. In addition, the common learning theories that its proponents use to advance its usage, and its continual evolution and implementation in today's classrooms will be discussed (Sections 2.4.3 and 2.4.4). Section 2.5 will examine the rich field of learning environments research. The origins of the field will be examined as well as its growth and development throughout the years (Section 2.5.1). Various ways of assessing learning environments will be discussed (Section 2.5.2) as well as an overview of the process required to develop learning environment instruments. Section 2.5.3 will review some of the most common learning environment instruments from a historical perspective, with special emphasis on the learning environment survey used in this study (CUCEI). Section 2.5.4 will discuss the related field of evaluating student attitudes toward science, again with emphasis on the instrument used in this study (TOSRA). Lastly, Section 2.6 will summarize the chapter with concluding remarks.

2.2 Conceptual Framework to Study Design

The conceptual framework of a study consists of the ideas and beliefs regarding the key concepts/variables/phenomena being studied and the relationships between them (Maxwell, 2013). The framework design was based on 1) existing theory and research, 2) personal experience, and 3) thought experiments combining 1 and 2 to create the theoretical model below. Each component of this conceptual framework will be discussed in the following sections.

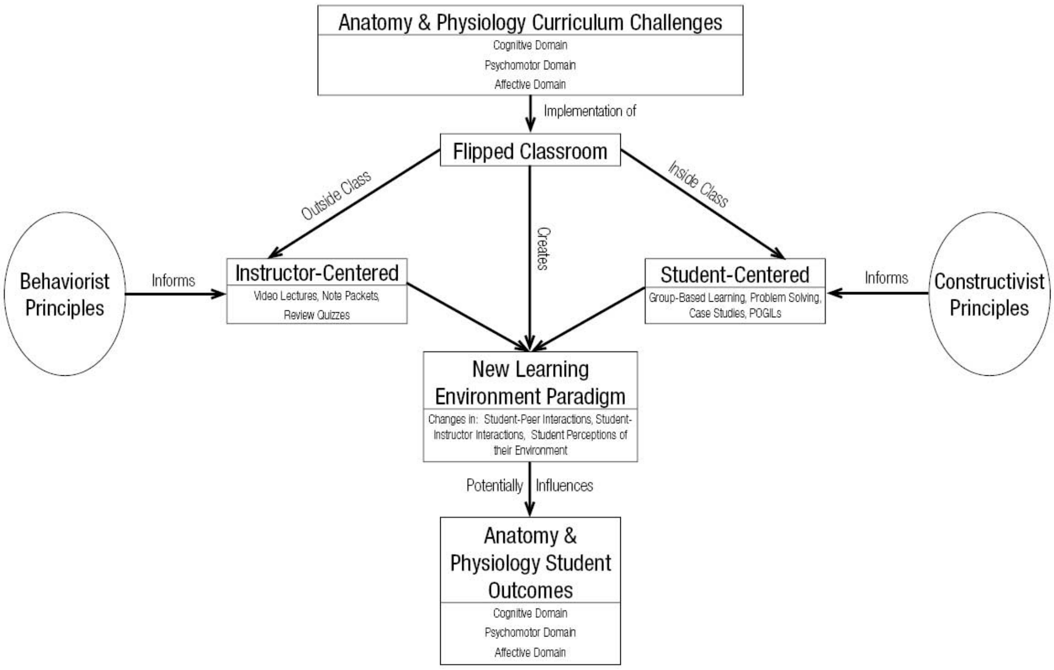


Figure 2-1 Conceptual Framework Concept Map

2.2.1 Anatomy and Physiology Curriculum Challenges

As introduced in Section 1.2 and detailed below in Section 2.3, undergraduate anatomy and physiology students face similar challenges as other classes. However, it is the unique structure of the curriculum (in volume, complexity and application to clinical thinking) which makes anatomy and physiology particularly challenging to students in all three domains of learning (i.e., the cognitive, psychomotor and affective domains, see Figure 2-2).

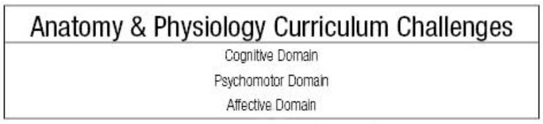


Figure 2-2 Anatomy & Physiology Curriculum Challenges

2.2.2 Flipped Classroom

While a detailed review of flipped classroom format literature will be discussed in Section 2.4, this section will discuss some basic principles as it relates to the conceptual framework of the study (Figure 2-3).

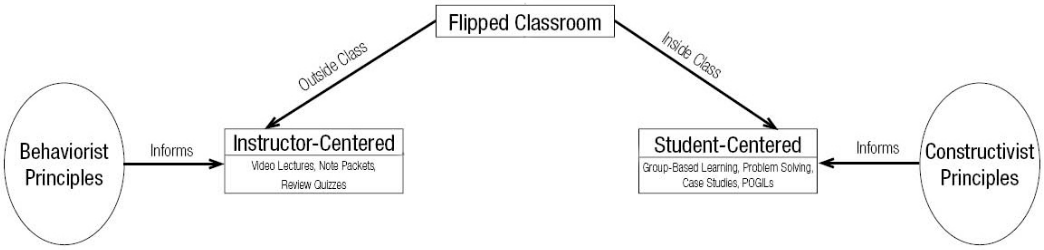


Figure 2-3 Flipped Classroom Basic Structure

2.2.2.1 Outside Class (Instructor-Centered/ Behaviorist Principles)

In many traditional undergraduate STEM classrooms, the structure is based on behaviorist principles, whereby an instructor-centered lecture is the primary form of transmission of information and acquisition of new knowledge by the student is basically a passive process (Khalil & Elkhider, 2016; Mattheis & Jensen, 2014). The structure of the flipped classroom moves this initial introduction of information from the classroom lecture to outside class, typically in a video lecture format. In addition, for the purposes of this study, note packets were provided which coincided with the lectures, and quizzes embedded within the videos served as formative assessments of the topics covered. In this way, the activities outside the classroom would mimic those within the traditional classroom (the control). This is explained in greater detail in Sections 2.4.3.1 and 3.5.1.

2.2.2.2 Inside Class (Student-Centered/ Constructivist Principles)

The initial transference of information having been done outside of class, the instructor can now utilize a more student-centered approach in the classroom, based on constructivist principles whereby the student must take an active role in the learning process. Having already been introduced to the material, the flipped classroom allows the students to engage with the material with higher-order thinking (Khalil & Elkhider, 2016; Mattheis & Jensen, 2014). While there are different strategies to engage with students in this manner, the primary ones used in this research were group-based and problem-solving activities, anatomy and physiology case studies, POGILs, and hands-on activities that would normally be only be available in a laboratory setting. This is explained in greater detail in Sections 2.4.3.2 and 3.5.2.

2.2.3 New Learning Environment Paradigm

Due to the fact that the flipped classroom is a relatively new instructional technique, this will be an entirely new learning environment for the vast majority of students. Hence, they will need to have a paradigm shift to their approach to learning in the classroom. This includes changing how they engage with the information, how they interact with the instructor, and how they interact with their fellow students (Figure 2-4). This is not only due to the changes in the classroom structure, where active participation by the student is required, but also because they must initially engage with the information prior to class in a new way.

Figure 2-4 New Learning Environment Created by Flipped Classrooms

2.2.4 Anatomy and Physiology Student Outcomes

The flipped classroom approach is intended to allow students to engage with the material in a deeper, more meaningful way. It does so by allowing for more student-to-student and student-instructor interactions (as opposed to the one-way transference of information in a traditional lecture). It also allows more time for hands-on activities that would normally be only available in a laboratory setting. In this way, the flipped classroom should influence student outcomes in the cognitive, psychomotor and affective domains (Figure 2-5).

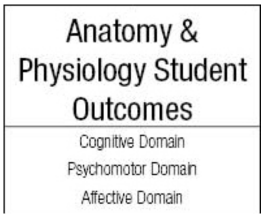


Figure 2-5 Potential Changes in Anatomy and Physiology Student Outcomes due the Implementation of a Flipped Classroom

2.3 The Undergraduate Anatomy and Physiology Classroom

As introduced in Section 1.2, this section will expand upon the nature of an anatomy and physiology classroom, it's foundation for future coursework and clinical practice, and the common difficulties endured by both students and instructors. Ultimately, a common theme found in much of the research is a search for the best pedagogical/andragogical methods to ensure comprehension, overcome misconceptions, and enhance critical reasoning in clinical settings.

2.3.1 Importance of Undergraduate Anatomy and Physiology

Study of early recorded history shows that there has been a long interest in the human body and its inner workings. From the early physicians in Egypt and Mesopotamia over 3,000 years ago to Hippocrates (c. 460 – c. 375 BC) "the father of medicine" and Aristotle (384 – 322 BC) in ancient Greece, we find the origins of the science we know today (Martini, Nath, & Bartholomew, 2011). The term we use to study the structure of the human body, "anatomy" is derived from the ancient Greek word *anatome* meaning dissection (*ana*- meaning "up" and *-tomy* "to cut"). Likewise, "physiology" the study of the function of living organisms, originates from the ancient Greek word *physis*, meaning "nature" and the suffix -*logia* meaning "study of" (Martini et al., 2011). And while it can be argued that medical science has advanced more in the last 50 years than the previous 2500, the importance of understanding how structure (anatomy) dictates function (physiology) and how function reflects structure, known as the principle of complementarity, has been acknowledged for centuries (Saladin, 2018).

Whether undergraduate Human Anatomy and Physiology (commonly abbreviated as HAP or A&P) are studied separately (i.e., one semester of Human Anatomy and one semester of Human Physiology), or combined (i.e., A&P I and A&P II), the complementarity of anatomy and physiology is always stressed (McKinley, O'Loughlin, & Bidle, 2019; Saladin, 2018). It is this understanding of the integration of structure and function which creates the foundation of basic biomedical knowledge. Woods et al. (2007) point out that one must have this basic biomedical knowledge if one is to have *clinical knowledge*, which they define as the knowledge of the signs, symptoms, and associated findings of disease. A&P introduces topics such as the description of disease processes, clinical procedures and health and fitness issues which are examined in greater detail in later, more advanced clinical coursework (Cliff & Wright, 1996). As such, A&P is a common prerequisite for pre-professional and professional health programs such as nursing, pre-medical and medical, exercise science, biochemistry, pre-dental and dental, and radiological technology (Gultice et al., 2015; Raftery, 2007; Stetzik et al., 2015; Vitali et al., 2020). For example, Raferty (2006) points out that this coursework provides the basis for the physical examination of patients, the interpretation of medical imaging, and the safe conduct of basic procedures in elective and emergency situations (Raftery, 2007). Multiple authors stress the importance of the A&P curriculum. For example:

> "A&P provides a foundation for nursing care, and thus successful engagement with the basics of anatomy and physiology is fundamental to the development of sound nursing skills and clinically-based critical thinking" (Johnston et al., 2015, p. 415).

> "The heavy penalty of not concentrating on sufficient anatomy education will inevitably lead to incompetent anatomists and healthcare professionals, leaving patients to face dire repercussions" (Sugand, Abrahams, & Khurana, 2010, p. 83).

> "Medical students need to acquire core anatomical knowledge to build a strong foundation for future clinical encounters and professional practice" (Singh, Bharatha, Sa, Adams, & Majumder, 2019, p. 1)

> "Understanding of the human body anatomy and physiology is pivotally important for preparing students of nursing for competent clinical practice" (El-Sayed & El-Sayed, 2013, p. 215).

> "Human anatomy and physiology is a foundational course for advanced topics in the basic or clinical sciences that put increasing emphasis on problem solving" (Cliff & Wright, 1996, p. S19).

In addition to the importance to clinical practice, success in A&P courses has also been shown to correlate with overall university Grade Point Averages (GPA) for both undergraduate health science and medical school students (Vitali et al., 2020). Several authors have noted that the academic skills necessary to be successful in A&P have the potential to aid students through their entire undergraduate journey (S. J. Brown et al., 2017; Pandey & Zimitat, 2007; Vitali et al., 2020). Likewise, others have observed that those who struggle in A&P often have difficulties in advanced clinical coursework (El-Sayed & El-Sayed, 2013). Brown et al. (2017) suggest that due to the ability of A&P courses to predict success in later study, academic achievement in A&P courses may be used to inform university policy regarding student progression.

2.3.2 Common Traits and Difficulties

While A&P is considered the foundation of the medical sciences, it is common for students to perceive it to be an arduous and challenging subject in their education (Singh et al., 2019) and many instructors and researchers find the current state of A&P education to be on a decline as student are not performing as well as in the past (Eseonu et al., 2013; Raftery, 2007). This can be attributed to several factors, including the unique nature of the content itself, lack of time spent on the topic, inadequate learning strategies utilized by the students, lack of proper educational resources, and A&P instructors who are experts in their field but not in instruction.

2.3.2.1 Unique nature of A&P Content

Obviously, there are many variables that contribute to student success, and every student is unique, bringing their own distinct backgrounds, perspectives, and learning styles. However, there are some commonalities within human anatomy and physiology classrooms that make it especially difficult for instructors to guide students toward success. Most prevalent, anatomy and physiology instructors are faced with the problem of how to help students to learn and retain a very large volume of complex information (Dobson, 2013). Significant effort is required by students to identify complex structures, their internal organization, and their relationships with other structures in the body (Vitali et al., 2020). This often leads students to be intimidated by what is often seen as one of the most content-dense and conceptually challenging courses within their curriculum. Furthermore, if students do not have a strong science background and have had a negative previous experience in the biosciences, they can become "science-phobic." Therefore, for students such as this, the challenging areas of human anatomy and physiology need to be presented in a manner that helps students overcome anxiety and support repetition/memorization learning and the conceptual understanding required (Johnston et al., 2015).

The difficulty of human anatomy and physiology within a biomedical curriculum is evident when one examines the high dropout, withdrawal and failure rates (Higgins-Opitz & Tufts, 2015; Slominski, Grindberg, & Momsen, 2019; Vitali et al., 2020). These high failure rates can often delay student entry into career pathways. In addition, it can decrease the diversity of the pool of applicants to professional programs, lower morale for both students and instructors, and increase the overall costs for students and the institution. (Gultice et al., 2015). One of the more challenging aspects of A&P as compared to other STEM courses is the different study skills necessary to be successful in anatomy versus physiology. While the complementarity of anatomy and physiology must always be at the forefront of instruction, showing the interdependence of the two subjects, it has been recognized that the subjects independently have a unique set of challenges for students. The study of physiology tends to be more conceptually challenging, requiring higher-order thinking to understand the dynamic interactions within the body, and research suggests is more difficult to learn than anatomy (Sturges & Maurer, 2013). The subject of anatomy is memorization heavy, requiring students to recall a large amount from a massive body of facts, and more importantly, to recall the pertinent facts in a clinical setting (Cliff & Wright, 1996).

2.3.2.1.1 Criticism of modern A&P: Students not prepared for clinical settings

A major concern stated by multiple authors is that modern health care professionals do not have sufficient anatomical knowledge. Signh et. al. state that "anatomy teaching in the medical and other health professional education programmes is on the decline and 'has fallen below a safe level' in recent years" (2019, p. 1). Others have stated that the teaching and learning of anatomy at both the undergraduate and graduate levels is in crisis, and point to the increasing surgical errors and the resulting increases in medical litigation, as evidence that there is a decline in the knowledge of professionals (J. P. Collins, 2009; Smith, Martinez-Álvarez, & Mchanwell, 2014). One of the first to document concern for the state of anatomy instruction was Older (2004) in the journal, *The Surgeon*. He writes that the "reduction in undergraduate teaching and knowledge of anatomy has caused great concern, not only for undergraduates, but also to postgraduate students, especially in surgery...... has set up a system that is allowing young men and women with poor knowledge of anatomy to become surgeons" (2004, p. 79). This criticism is acknowledged by the students themselves. Woods et. al. (2007) found that only 14% of final year medical students felt confident in their anatomical knowledge. Smith and Mathias (2011) found that the majority of new doctors (68.3%) were concerned that there was still so much anatomy they did not know and that they may lack the ability to properly apply the knowledge they did have in practice. Fifty-four point six percent claimed to forget much of the anatomy they had learned (Smith & Mathias, 2011).

2.3.2.1.2 Evolution of Anatomy Instruction

Because of its importance in medical education, traditional anatomy instruction consisted of substantial study hours through both undergraduate and graduate programs with extensive examination (Wilhelmsson et al., 2010). However, several authors have noted the time allotted for A&P has been dramatically reduced in recent decades because of the requirements to accommodate other subjects in an ever-expanding biomedical curriculum. (J. P. Collins, 2009; T. J. Collins, Given, Hulsebosch, & Miller, 1994; Eseonu et al., 2013; Johnston et al., 2015; Older, 2004; Priyadharshini et al., 2019; Smith et al., 2014; Vitali et al., 2020). Specifically, Older (2004) points out that time and content of gross anatomical instruction have decreased more than 50% compared to 25 years ago. Furthermore, major criticisms have been levelled against the traditional structure. Many considered it to be an overcrowded curriculum, overemphasizing clinically unconnected facts and memorization, with didactic lectures which were teacher-centric. (Older, 2004). This led to educational reforms shifting medical education away from unconnected facts and excessive memorization toward one of clinical understanding. While this shift is welcomed by most, in some cases it can be argued that the pendulum has swung too far the other way, with little to no cadaver-based gross anatomy being taught. Raftery goes onto state that "the anatomy learnt in the undergraduate course is so cursory that a lot of new learning is required and very little recall" (2007, p. 2).

2.3.2.1.3 Evolution of educational resources for A&P

Traditionally, human cadavers are often seen to be the optimal way to illustrate certain anatomical and physiological concepts. Raftery summarizes Professor Harold Ellis's stated benefits of cadaver dissection (Ellis, 2001; Raftery, 2007):

- teaches the basic language of medicine
- teaches some manual dexterity
- introduces an understanding of three-dimensional anatomy
- introduces the concept of biological variation
- introduces some common pathological conditions
- introduces students to team working
- introduces communication skills between individual students and groups
- acclimatizes students to the reality of death and teaches respect for the body
- is a prime example of 'self-directed learning'

Older (2004) agrees that there are many benefits to cadaver dissection, including the fact that the labs create an environment that stimulates a social bond between the participants which would stimulate the affective domain of learning.

Multiple studies support the idea that dissection of the whole human structure provides the valuable format for learning anatomy, especially the integrated functions of multiple structures to maintain homeostasis (Griff, 2016; Smith & Mathias, 2011). One study found that 77% of new medical school graduates said that using human cadaveric prosections (a dissection by an experienced anatomist) was an effective was to learn anatomy (Smith & Mathias, 2011). However, human cadavers are not a viable option for many schools due to the lack of donated bodies, cost of maintenance, space requirements, governmental regulations, safety concerns, and regions where the use of human cadavers are prohibited (Anderton, Shan Chiu, et al., 2016; El-Sayed & El-Sayed, 2013; Johnston et al., 2015; Sugand et al., 2010). There have been attempts to modernize the study of A&P and enhance or even replace dissection with other modalities. Examples of other modalities are interactive multimedia, procedural anatomy, surface and clinical anatomy, plastic models, and imaging (Sugand et al., 2010). However, many still believe that these other methods cannot lead students to the reasoning necessary for clinical knowledge the same way investigative dissection of real tissue can (Miller et al., 2002).

2.3.2.2 Inadequate learning strategies by A&P students

In addition to the inherent difficulties due to the nature and structure of an A&P class, students often approach the course with improper learning strategies. Some simply enter the course lacking the basic math, reading, or study skills necessary to succeed (Gultice et al., 2015). For many students, unsuccessful learning strategies makes A&P feel overwhelming. Many report that they feel there is simply too much information. It has also been noted that many students feel there is often a disconnect in the material, in that they cannot always articulate the relevance of a topic or understand why a topic was important. In addition, many struggle with the 3-dimensional component to A&P, reporting difficulty visualizing orientation and judging scale (Smith et al., 2014).

When examining student learning strategies in A&P, many researchers take the classic view of student learning, that students tend to take either deep, strategic or surface approaches to memorizing the large volume of material within the course (Smith & Mathias, 2011; Wilhelmsson, Dahlgren, Hult, & Josephson, 2011). A surface approach is seen as one of rote memorization and simple reproduction of facts. It is seen as a very common tactic within an A&P classroom. A deep approach is characterized by a motivation to understand the topic and make connections. A strategic approach is focused on assessment, in that students adopt whatever learning strategy they perceive will achieve the best test results. Many have pointed out that undergraduate students need to be encouraged to move away from surface learning approaches and toward deeper learning strategies, which are characterized by a drive to understand underlying principles and concepts by grappling meaningfully with content (McLean et al., 2016). Specifically, with anatomy, knowledge must be constructed into meaningful and useful components and rote memorization should be replaced by investigative or active learning approaches (Griff, 2016; Smith & Mathias, 2011). It should be pointed out to students that anatomists do not depend on rote memorization, but think about the how and why, naturally integrating form and function (Griff, 2016; Miller et al., 2002). However, while researchers stress that most anatomy students have a deficit in deep learning, specifically in contextualizing and creating meaningfulness within the information, it has also been acknowledged that surface learning can still have its place in the initial stages of learning and a balance must be sought between the two (Smith et al., 2014; Wilhelmsson et al., 2010).

It is not uncommon for those who have taken A&P to not retain much information. While this was notated previously by the number of medical school graduates who felt there was still much anatomy they did not know, this is also seen at the undergraduate level. Griff (2016) points to a retention study of allied health students that found that students who completed an elementary physiology course did not perform statistically different in an advanced physiology course than those who had no prior physiology experience (Griff, 2016; Richardson, 2000).

2.3.2.3 Instructor pedagogy

Often A&P courses, and biology courses in general, are taught by scientists who have good content knowledge, but have very little formal education in the science of instruction (Mattheis & Jensen, 2014). It is not uncommon to find biology departments with teaching staff with greater or lesser backgrounds in the biosciences and quite variable teaching skills (Johnston et al., 2015). So, while their position requires them to be involved in designing, developing and delivering curriculum, most lack any knowledge in the scientific rationale for proper instructional design. The goal of proper instructional design is to increase instructional efficiency and facilitate student learning (Khalil & Elkhider, 2016). Unfortunately, this is often overlooked, and many instructors simply teach as they were taught. They do so in a teacher-centered, lecture-based format, while having little-to-no knowledge of evidence-based, effective teaching techniques. The teacher-centric format has been found to prioritize memorization of facts over conceptual understanding and higher-order thinking (Mattheis & Jensen, 2014; Stetzik et al., 2015). Many instructors acknowledge being challenged by the need to learn new approaches to teaching (Megaw & Zimanyi, 2019). However, as Miller et al. point out, "We have come to recognize that what and how we teach in A&P does not necessarily correspond with the content and skills that we think are most important for our students" (2002, p. 74). Somewhat ironically, Miller et al. also points out that "Traditional A&P textbooks match the way we teach now, not the way we would like to teach" (2002, p. 74).

What can make instruction harder for A&P professors is the student diversity that is common to the class. Because of the many majors where A&P is a requirement, it is commonly a large class with a diverse group of students with varied backgrounds. The diversity can come in many ways, including academic ability, learning styles, ethnicity, varying socioeconomic and educational backgrounds, and overall preparedness for university life and study. All of which can have an impact on learning success in A&P and must be taken into consideration by the instructor (Higgins-Opitz & Tufts, 2015; Megaw & Zimanyi, 2019; Miller et al., 2002).

There has been a call for reforms in A&P instruction for quite some time. Griff (2016) cites successful instructional techniques such as scaffolding of new information on what students already know, interweaving easy and challenging material, the use of repetition, and frequent formative assessments. Collins (2009) suggests that "anatomy should be taught, learned and assessed within a clinically meaningful context, which helps the learner to integrate an understanding of normal function with the recognition of normal structure" (2009, p. 19). Miller et al. (2002) stress the importance of creating the proper instructional climate, one which focuses on the process of discovery. They suggest that it should be presented in such a way that emphasizes problem solving and application into clinical care. In the most basic sense, Mattheis and Jensen (2014) summarize it best in that they call for a "shift from teacher-centered to student-centered learning" (p. 321). This is precisely the goal of the flipped classroom approach.

2.4 Flipped Classrooms

As introduced in Section 1.2.3, a flipped (also referred to hybrid or inverted) classroom approach uses technology to introduce content to the students prior to a scheduled class, thereby allowing students to engage with the content, often in a more applied manner, and at a deeper level within the classroom (Strayer, 2012). Or, in a sense, what is normally done in class is done as homework and what is normally done as homework is done in class, hence the term "flipped" (Bergmann & Sams, 2012; Herreid, C. F., & Schiller, 2013; Lage et al., 2000). However, in practice, the flipped classroom is more than a simple re-ordering of classroom and homework activities. Rather, careful consideration must take place to properly align inside-class and outside-class materials and to continually collect formative feedback of students throughout the process (O'Shea, 2020a). In-class activities are to be group-based and student-centric interactive learning activities (Bishop & Verleger, 2013). In this way, the flipped classroom provides students with a student-centered learning context, changing the role of the lecturer to one of a "learning guide" to address student questions when needed (Peter, Khoo, Scott, & Round, 2016). The information presented prior to class may be presented in multiple forms, such as notes with images, augmented reality, and video lectures (the most common form) (Bishop & Verleger, 2013; Ferrer-Torregrosa et al., 2016). Hence, this "approach repurposes class time to focus on application and discussion; the acquisition of basic concepts and principles is done on the students' own time before class" (McLean et al., 2016, p. 47). This would seem to be an ideal solution for an anatomy and physiology classroom because many anatomy educators agree that success depends on overcoming the student misconception that they must focus solely on memorization and instead promote a focus on reasoning skills to understand anatomy and physiology (Miller et al., 2002).

2.4.1 History of the Flipped Classroom Approach

Over the past several decades, there have been several calls to reform science education at all levels (K-12 to postgraduate). The role of students has often been seen as a passive note-taker, simply repeating what has been presented to them (Michael, 2006). Often termed "bulimic learning" as students continually resort to memorizing the material and regurgitate the information on examinations (Rotellar, C., & Cain, 2016). These critiques have urged educators to adopt new approaches that "actively involve the student in the learning process, that focus on problem-solving as well as memorization" (Michael, 2006, p. 159). Research from several related disciplines (e.g., learning sciences, cognitive psychology, educational psychology) suggested a student-centered, active learning pedagogy (i.e., an instructional method that involves students in the learning process) is what is needed to engage students to appreciate the scope, meaning, and limitations of science (Bishop & Verleger, 2013; Michael, 2006). During this same period, there has been increased availability of information via internet and computer applications. Professionally produced science video lectures became much more accessible to students at home (e.g., MIT OpenCourseWare, Khan Academy, Bozman's Science, etc.). Also, software tools have been made available for educators to create their own online lectures (e.g., Camtasia, Adobe Presenter, PaperShow, etc.) as well as tools to deliver the information online to their students (e.g., YouTube, Blackboard, Moodle, etc.). The combination of educational research, the calls for education reform, and accessibility of video lectures helped lead to the development of the flipped classroom approach (Bishop & Verleger, 2013; Herreid, C. F., & Schiller, 2013; Michael, 2006).

The origins of the term "flipped classroom" as well as its popularity are often associated with Johnathan Bergmann and Aaron Sams, two high school chemistry teachers from Woodland Park, Colorado, USA (Bergmann & Sams, 2012; Walker, Tan, Klimplová, & Bicen, 2020). Although both Bergmann and Sams acknowledge that they did not originate the concept nor the term "flipped classroom", their initial success spawned excitement for the process. Originally, they began recording their lectures for students who missed class (their high school was in a rural area and students in extracurricular activities tended to miss a lot of school while traveling on buses to various competitions). After posting their lectures online and getting good feedback from their students as well as other students and instructors from all over the world, it occurred to Sams that, "The time when students really need me physically present is when they get stuck and need my individual help. They don't need me there in the room with them to yak at them and give them content; they can receive content on their own" (Bergmann & Sams, 2012, pp. 4–5). This lead to their initial implementation of the flipped classroom format and through trial and error, developing a system which proved to be very successful with their students (Bergmann & Sams, 2012).

However, prior to the excitement of the flipped classroom approach generated by Bergmann and Sams, other instructors were experimenting with the concept. One of the first papers published on the subject is "Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment" by Lage, Platt and Treglia (2000). The researchers used a variety of teaching styles in class to appeal to the multiple learning styles of their economic students at the University of Miami, Ohio. Outside of class, students watched lectures on VHS tapes or PowerPoints with audio. They reported generally positive feedback from both students and instructors (Lage et al., 2000; Megaw & Zimanyi, 2019). Other early successes followed, such as at the University of Wisconsin-Madison where a computer science course for engineers was flipped. Students there gave significantly higher ratings to all aspects of the course (Foertsch, Moses, Strikwerda, & Litzkow, 2002). However, after those early reported successes, later studies have been decidedly mixed, with some students reporting positive impacts on learning and others perceiving a lack of structural support. It quickly became apparent that the flipped classroom approach is not simply putting lectures online and doing "homework" in the classroom. Flipped classrooms must use specific design principles to increase interaction between students and teachers in a meaningful way (Rotellar, C., & Cain, 2016).

2.4.2 Structuring a Flipped Classroom

While the flipped classroom method has shown a lot of promise, the creation and implementation of a flipped classroom can be challenging for both teacher and student. As Walker et al. (2020) point out, "As with all disruptive paradigms, the learning curve for flipping the classroom can be steep for students and faculty" (p. 10). Arner (2020) suggests that the following questions must be addressed when designing a flipped classroom:

- What do students do in the pre-meeting? In class?
- How do teachers create or curate high-quality learning activities?
- What evidence-based practices can teachers incorporate in their flipped classrooms?

- How do teachers know if their students are learning?

As notated previously, there is no strict design standard for creating a flipped classroom. However, Kim et. al. (2014) proposed nine design principles for instructors to consider. The first three were adopted and validated from design suggestions by Brame (2013): 1) Provide an opportunity for students to gain first exposure prior to class; 2) provide an incentive for students to prepare for class; 3) provide a mechanism to assess student understanding. The next six principles were developed out of research by Kim et. al. to create better student-centered learning: 4) Provide clear connections between in-class and out-of-class activities; 5) Provide clearly defined and well-structured guidance; 6) Provide enough time for students to carry out the assignments; 7) Provide facilitation for building a learning community; 8) Provide prompt/adaptive feedback on individual or group works; 9) Provide technologies familiar and easy to use (2014, pp. 43–46). As with most instructional formats, the specific curriculum for outside-class and inside-class activities is left up to the instructor's best judgement, however these principles can be considered "best practices" in the creation of a flipped classroom. It is considered essential that instructors carefully consider the design of the flipped classroom and keep the student perspective at the forefront. As stated by O'Shea (2020), "It is unlikely that flipped learning will fully reach its potential as an educational approach if the student experience is not at the center of the material development process" (O'Shea, 2020a, p. 290)

2.4.3 Learning Theories Utilized within the Flipped Classroom

For most researchers, the flipped classroom is seen as a unique combination of two learning theories once thought to be incompatible, constructivism and behaviorism. In-class activities consisting of active, problem-based learning founded on constructivist principles and while outside-class lectures of direct instruction are founded on behaviorist principles (Bishop & Verleger, 2013; Limniou, Schermbrucker, & Lyons, 2018). For many advocates of the method, it is this unique marrying of didactic education based on behaviorism and clinical practice performance based on constructivism which gives the flipped classroom method an advantage over either individual instructional design method (Walker et al., 2020).

This section will review both behaviorist and constructivist principles as well as other learning theories commonly associated with the flipped classroom.

Behaviorism, constructivism and humanism are commonly seen as the three dominant psychological schools of thought which are relevant to learning theory. It should be noted that while each represents a differing perspective on learning, they can, in fact, complement each other. Other learning theories and theorists are often categorized as under one of these three theories, (Aubrey & Riley, 2019; Khalil & Elkhider, 2016). While this section will focus on behaviorist and constructivist principles, humanist principles, which focus on the needs of the individual learner, can be applied to the flipped classroom as well. Instruction based on humanist principles focus on what is needed for the personal and emotional growth of the student. In the flipped classroom this can be achieved inside the classroom where instructors may have more time to meet the needs of the individual learner than in a traditional classroom (Aubrey & Riley, 2019; Bishop & Verleger, 2013).

2.4.3.1 Outside Class - Behaviorist Principles

In behaviorism, learning is considered to be the acquisition of a new behavior, and the learner's role in the learning process is passive (Khalil & Elkhider, 2016). For behaviorists, behaviors are predictable and measurable, and are acquired through conditioning. So, from a behaviorist perspective, learning is simply a matter of stimulus and response. Behaviorist principles are most often traced back to the works of Ivan Pavlov, Edward Thorndike, John Watson, and B. F. Skinner. Operant conditioning was the term coined by Skinner to describe how learned behaviors are created by associating it with a consequence. In the classroom setting, operant conditioning can been seen by an instructor providing a stimulus, the student responding to the stimulus, and then the instructor provides either positive or negative reinforcement based on the student's response (Aubrey & Riley, 2019; Knowles, Holton III, & Swanson, 2015). Common examples of teaching methods utilizing a behaviorist approach are lectures, stimulations, demonstrations and programmed instruction (Khalil & Elkhider, 2016).

A traditional lecture at the university level is commonly associated with behaviorist principles. The top-down approach, where the professor is seen as the authority and archetypal "Sage on the Stage", relaying whatever information is on the agenda down upon the students. This traditional format is often what is expected by students. Studies have shown that undergraduate students perceived education to be a hierarchical experience, with scholarly instructors passing on information. Somewhat ironically, the same studies showed instructor's perceptions of their experiences with students to be more collaborative in nature (O'Shea, 2020b). In the most common form of a flipped classroom, this behaviorist lecture format is done outside of class in the form of a video. These video lectures allow the student to have the background knowledge to properly participate in deeper discussions and activities during class time. Student success is dependent on their level of participation prior to class, as instructors structure in-class activities based on students being properly prepared (Walker et al., 2020).

Pre-recorded video lectures are considered to be a good medium for outside-class lessons because it links audio and visual together, creating a multisensory experience for the learning. According to El-Sayed and El-Sayed, "Video-based lectures provide a unique opportunity to present, teach, and internalize information; they are also excellent venues for focusing the students' attention on specific details based on the prepared material itself" (2013, p. 218). While there are several variables that will contribute to a successful video lecture, one most often cited is video length. While research is inconclusive as to the best format for video production, it has been noted in several studies that students prefer shorter video lectures. Videos between 3 and 12 minutes have shown to be effective in the flipped classroom format (Walker et al., 2020). A key feature that is often incorporated into the video is some sort of evidence of the student's interaction with the pre-class content. This is often accomplished by including an interactive component to the video (e.g., incorporating quiz questions). Multiple studies have found the inclusion of an interactive component to lecture videos to be highly effective for learning (Arner, 2020).

2.4.3.2 Inside Class – Constructivist Principles

From a constructivist perspective, learning is not something students passively acquire while their teacher lectures. Rather, students are active participants in the learning process, constructing understanding and meaning by building upon knowledge they already have (Aubrey & Riley, 2019; Khalil & Elkhider, 2016). Constructivism is often further divided into two forms: cognitive constructivism and social constructivism. In cognitive constructivism, individuals make personal meaning of their learning experiences, with new knowledge analyzed with what they already know. Cognitive constructivism is typically associated with John Dewey, Jerome Bruner and most notably, Jean Piaget (Aubrey & Riley, 2019; Knowles et al., 2015). Social constructivism builds upon this concept by incorporating the aspect of social interaction. The interaction of individuals in the learning environment, language and discourse, cultural and social backgrounds of the learners are all emphasized in the construction of learning. Social constructivism is typically associated with Lev Vygotsky (Aubrey & Riley, 2019).

Piaget considered himself not so much a developmental psychologist, but rather a "genetic epistemologist" (someone who studies the origins of knowledge) (Young, 2011a). Perhaps most pertinent to the flipped classroom approach is his *Theory of Cognitive Conflict* (Bishop & Verleger, 2013). Cognitive conflict is defined as, "a psychological state involving a discrepancy between cognitive structures and experience, or between various cognitive structures (i.e., mental representations that organize knowledge, beliefs, values, motives, and needs). This discrepancy occurs when simultaneously active, mutually incompatible representations compete for a single response. The detection of cognitive conflict is thought to trigger compensatory adjustments in executive control processes, which serve to reduce and prevent subsequent instances of similar cognitive conflict" (Waxer & Morton, 2012). According to Piaget, the mental structures an individual creates from the information around them is known as a *schema*. Experiences can add and grow these representations (schemas) so one can make sense of the world. If one is able to use their existing schemas with the information taken in from the senses to make sense of the world, they are said to be in *equilibrium*. The adaptation of an existing schema in order to make sense of a new experience is known as *assimilation*. If a new experience will not fit into an existing schema, a new schema will be produced. This is known as *accommodation*. In this way, individuals construct new knowledge based on experiences and existing knowledge (Aubrey & Riley, 2019; Young, 2011b).

One of the major critiques of Piaget was related to his assertion that learning happens best in isolation. Lev Vygotsky, who was a contemporary of Piaget's, agreed that knowledge was constructed (Aubrey & Riley, 2019). However, Vygotsky saw learning as a result of guided social interactions. In his view, cognitive development could only be understood if you take into account the social and cultural context of the learner. Vygotsky defined the term to describe the gap between the developmental level of the learner and the level they could reach with the right educational support and guidance as the *Zone of Proximal Development*. The assistance from an adult or more capable peer to help a learner solve problems and decrease the gap is known as *scaffolding* (Aubrey & Riley, 2019; Bishop & Verleger, 2013; Podolskiy, 2012).

Bishop and Verleger (2013) make the case that the constructive rationale for student-centered learning for the in-class activities in a flipped classroom are derived initially from the works of Piaget and Vygotsky. They show how other learning theories and theorists are traced back to these two pioneers. Specifically, Bishop and Verleger cite other authors who point out that the theories of cooperative learning, collaborative learning, peer-assisted learning, problem-based learning, experiential learning and learning styles can connect in some way to Piaget and Vygotsky (Bishop & Verleger, 2013). At the core of all these theories is the concept of active learning. According to Michael (2006), active learning is defined as "The process of having students engage in some activity that forces them to reflect upon ideas and how they are using those ideas. Requiring students to regularly access their own degree of understanding and skill at handling concepts or problems in a particular discipline; the attainment of knowledge by participating or contributing. The process of keeping students mentally, and often physically, active in their learning through activities that involve them in gathering information, thinking, and problem solving" (p. 160). Michael (2006) points to a growing body of evidence to support active learning pedagogies. That said, the shift from a lecture-based format to an active learning approach changes the way students interact with each other, the material and the instructor. Hence, the flipped classroom transforms the learning environment for the student as well as the instructor and has the potential to change students' attitudes to learning anatomy and physiology (and science in general).

2.4.3.3 Other contributing theories

While the theories of behaviorism and constructivism or their derivatives tend to dominate the discussion of rationale in flipped classroom research, there are several other learning theories that are mentioned in the literature regarding the structure and/or analysis of the flipped classroom.

2.4.3.3.1 Bloom's Taxonomy of the Cognitive Domain

In Bloom's seminal work, *Taxonomy of Educational Objectives, Book I Cognitive Domain* (1956), he created a hierarchical classification of reasoning skills. His goal, in part, would be that the six progressively more difficult levels of cognitive function could be used by instructors to guide curriculum and more reliably assess students. Table 2-1 summarizes the original six levels (they have been modified by others since), from simple to more complex (Aubrey & Riley, 2019).

Table 2-1

Bloom's Levels of the Cognitive Domain Adapted from Aubrey and Riley (2019)

Level	Description of Level
Knowledge	Facts, recall, categorization, theories and abstractions
Comprehension	Making sense of what things mean and how they relate to each other
Application	Applying knowledge to different situations
Analysis	Breaking down knowledge into its constituent parts to gain a clearer understanding of the whole
Synthesis	Bringing together the separate constituents to create a new whole
Evaluation	Reflecting on knowledge and making judgements

Bloom's cognitive domain taxonomy has been referenced in flipped classroom research regarding the development and analysis of the flipped classroom. The rationale being that in a flipped classroom, the lower cognitive levels (i.e., knowledge and comprehension) are to be emphasized outside class when introducing the material. The higher levels, which require more abstract thought, are emphasized in the in-class material where active learning is stressed. With this in mind, Gilboy et al. described how being strategic and deliberate in development of the flipped classroom, faculty can "achieve all levels of Bloom's taxonomy with the instructor present." (Gilboy, Heinerichs, & Pazzaglia, 2015, p. 109). To analyze a flipped classroom for first-year medical students, Morton and Colbert-Getz (2017) used Bloom's taxonomy when structuring the gross anatomy items on a Foundations of Medicine final examination. They showed that while there was no difference between performance and knowledge between the flipped students and students in a traditional lecture, the flipped students did perform better on analysis items (Morton & Colbert-Getz, 2017). Day (2018) had similar results analyzing a flipped gross anatomy classroom for physical therapy students. Using Bloom's taxonomy, Day showed that previously lower-performing students performed better on higher-level questions in the flipped classroom than a traditional lecture (Day, 2018).

2.4.3.3.2 Information Processing Model and Cognitive Load Theory

The dominant theory of learning in cognitive psychology is the information processing model (Khalil & Elkhider, 2016). This theoretical model of learning sees learning as a process of knowledge acquisition similar to information processing in computers (Hoidn, 2017). The model consists of three basic types of memory. The first, sensory memory, represents the input of information from the environment via the senses, with typically sight and sound being the most pertinent senses when it comes to learning new information in the classroom. Sensory memory is passed onto working memory when the individual is consciously paying attention to it. For information to be learned and moved into the third type of memory, long-term memory, it must first move into working memory. Working memory is limited, and is often considered to be a vital factor to take into account when designing instruction (Abeysekera & Dawson, 2015; Khalil & Elkhider, 2016). When students are processing information (i.e., attempting to move information from working memory to long-term memory), they will often use two types of rehearsal strategies: maintenance and elaborative rehearsal. Maintenance rehearsal is repeatedly going over the information but not organizing it into a mental structure. This is also known as rote memorization and, as noted previously, is very common in anatomy and physiology classrooms. Elaborative rehearsal is the process of organizing the information in working memory in order to reach meaning. This is also known as deep learning (Khalil & Elkhider, 2016).

Cognitive load theory is based upon the information processing model. It assumes that for effective learning to take place the instructional conditions should be aligned with the leaner's cognitive architecture. According to the theory, there are three types of cognitive load. Intrinsic cognitive load refers to the core of the subject focused on in working memory. Extraneous cognitive load refers to the information in working memory that does not contribute to the proper construction of meaning. Lastly, germane cognitive load refers to the information being processed in working memory that contributes to meaning (the production of schemas). Within instructional design using the cognitive load theory, the concept is to reduce extraneous cognitive load (those things that are unimportant and/or distracting), manage intrinsic cognitive load and promote germane cognitive load (Khalil & Elkhider, 2016; Sweller et al., 2011). It has been proposed that both aspects of the flipped classroom can be optimized according to the cognitive load theory. Outside the class, students can control rate of information input from pre-recorded lectures by pausing and rewinding. In this instance it has been proposed the flipped classroom environment can reduce cognitive load. Furthermore, with dynamics allowed for in-class design, instructors may be able to better tailor instruction to the expertise of the students, further enabling more appropriate management of cognitive load (Abeysekera & Dawson, 2015).

2.4.4 Research of the Flipped Classroom

Although a relatively new instructional technique, because of its popularity, research in the flipped classroom approach has been robust for the past decade. However, because of a lack of a strict definition and wide variation of implementation of the flipped classroom, the results of the research as to its effectiveness have been mixed (Abeysekera & Dawson, 2015; Bishop & Verleger, 2013; van Alten, Phielix, Janssen, & Kester, 2019). In addition, there is also great variability in how studies of the flipped classroom are conducted, including many small studies and others which do not have a control group, hence making generalization quite difficult. Ultimately, researchers are still interested in determining if flipped classrooms are beneficial and if so, what are the variables that make it so? This can be seen in a recent meta-analyses of the flipped classroom approach conducted by van Alten et al (2019). In their analysis, 114 studies across multiple disciplines and educational levels were examined. They noted that there was great diversity in the studies reviewed, both in implementation of the flipped classroom as well as how the studies were conducted. They found that students in the flipped classrooms tended to perform better than in traditional classrooms and were equally satisfied with their learning environment. They concluded that "The main implication following our results is that flipped classrooms are worth implementing. Careful attention should be paid, however, to the design of the flipped classroom as simply flipping before and during classroom activities might be not enough" (van Alten et al., 2019).

2.4.4.1 Assessing the Flipped Classroom Approach in Higher Education

Most early studies of the flipped classroom took place in K-12 settings. From the early studies, systematic reviews and meta-analyses, it appears that the practice of flipping the classroom did not truly take hold at the undergraduate level until late in the first decade of the 2000s (O'Shea, 2020b). O'Flaherty and Phillips (2015) did a scoping review of the research in flipped classrooms in higher education from the time period 1995-2014. Their goal was "to establish how key aspects of the flipped class contribute to its effectiveness and to an improved student learning experience" (O'Flaherty & Phillips, 2015, p. 86). A scoping review was chosen over a systematic review in that the latter are utilized to answer specific questions while the former is better for fields with limited rigorous evidence. This structure allows a scoping review to identify gaps in the existing evidence and make recommendations for future research. After initial screening, 28 articles were included in their review. From their review, they recommend that future research consider indicators of student engagement other than examination scores. In addition, they found an under-utilization of conceptual frameworks in designing flipped classrooms. This often led to a disconnect between the pre-class and in-class activities, lack of clarity for the students and a heavy content focus. Lastly, they also called for stronger evidence in measuring student outcomes in higher-order thinking (O'Flaherty & Phillips, 2015).

Lundin et al. (2018) expanded on O'Flaherty and Phillips' review, by conducting a systematic review of studies from 2000 to mid-2016. Their goal was "to investigate what constitutes research on flipped classrooms and in particular to examine the knowledge contributions within the field so far and relate them to the wider research topic of educational technology in relation to higher education" (Lundin et al., 2018, p. 3). Prior to the screening process, they examined the 530 most-cited articles from 2010 to 2015 on the flipped classroom approach. This was done in order to investigate the potential increasing interest in the method. They found only four noteworthy publications in 2010, two in 2011, then growing to 16 in 2012, and then to 76, 136, and 296 over the following three years. In addition to the growth in general interest in the method, they noted an increase in interest in the tertiary level, with only 38% of the total publications investigating flipped classrooms in higher education in 2012, but up to 73% in 2015. Thirty-one articles from the initial 530, which were cited 15 times or more, were chosen for the review. They found that there is often a disconnect between the models of instruction and the espoused learning theories in research designs. For future research, they called for a stronger alignment between educational science and subject-specific didactics so that results and research designs can be compared and developed (Lundin et al., 2018).

2.4.4.2 Assessing the Flipped Classroom Approach in Health and Medical Sciences

Within the health and medical sciences, there have been several studies examining the effectiveness of the flipped classroom approach as well as several systematic reviews. For example, Betihavas et al. (2016) conducted a systematic review of the flipped classroom implemented in nursing education. Twenty-one articles of primary research were initially reviewed for relevance with nine ultimately included based on inclusion/exclusion criteria. Each of the nine were further critically appraised utilizing the 11 quality indicators as established by Buckley et al. (2009). Five studies ultimately were found suitable for the review, two quantitative studies and three mixed methods. They found that flipped classrooms yielded neutral or positive academic outcomes and mixed results with student satisfaction. They reported their findings were similar to other reviews in other disciplines (Betihavas et al., 2016). Chen, Lui and Martinelli (2017) followed by doing a systematic review of flipped classrooms in medical education. As the topic of medical education is broader than just nursing education, Chen et al. were able collect more than Betihavas et al., with 46 studies included in the final analysis, across 17 distinct medical specialties. They concluded that the flipped classroom approach in medical education shows promise, specifically as it relates to increasing learners' motivation, task value and engagement, however results were mixed as it relates to changes in behavior, professional practice and patient outcomes (Chen et al., 2017). Evans et al. (2019) followed with another systematic review of flipping the classroom in health care higher education. Here, 24 studies met the inclusion criteria within the educational fields of chiropractic health (2), medical (11), nursing (5), pharmacology

(5), and public health (2). They found that “the academic outcomes of the flipped classroom approach for the education of health professions students did not reveal compelling evidence of the effectiveness of the method above that of traditional classroom approaches” (Evans et al., 2019, p. 77). However, they also acknowledged that many of the studies examined had methodological weaknesses and called for stronger experimental designs in the future. For this reason, they concluded much like Betihavas et al. that the flipped classroom approach “has either a positive or equal effect on academic outcomes as the traditional lecture method” (Evans et al., 2019, p. 77).

2.4.4.3 Assessing the Flipped Classroom Approach in Anatomy and/or Physiology Classrooms

While there has been a relatively large amount of research into the flipped classroom method, specifically studying the implementation of the method in a human anatomy and/or physiology class has been rare. In an early study, DeRuisseau (2016) implemented the flipped classroom approach in an undergraduate anatomy and physiology classroom in the 2014-2015 academic year. The examinations of the flipped model were compared to traditional lecture examinations from previous years. DeRuisseau compared letter grades in both formats with students’ major GPAs. It was found that 85.5% earned a higher grade than their major GPA in the flipped class while only 42.2% did the same in the traditional. The study also found that the flipped classroom approach allowed for more time for active learning and allowed for a transition of assessment questions to include higher-order thinking activities. DeRuisseau noted that along with the outcomes benefit, student evaluations did suffer and called for more detailed studies comparing the traditional and flipped formats. Ultimately, DeRuisseau concluded that “Although the flipped classroom offers its own challenges, multiple lines of evidence now support that the benefits outweigh the risks. Increased higher-order thinking activities and assessments along with more time to cover the most difficult concepts make it a rewarding endeavor for both students and instructors” (DeRuisseau, 2016, p. 527).

As mentioned previously, Morton and Colbert-Getz (2017) implemented the flipped classroom approach to the gross anatomy section of a class for first-year medical students. They found that the flipped classroom method improved students’ ability to perform on analysis items on the final examination, but there was no difference in their ability to recall or recognize material on the final examination. By categorizing items on the final examination by levels of cognition based on Bloom’s taxonomy, Morton and Colbert-Getz were able to show the distinct potential advantages of active learning in the flipped classroom, as well as a potential reason why some studies did not find any significant differences in performance (i.e., because the assessments in the study did not properly assess the higher cognitive levels) (Morton & Colbert-Getz, 2017).

Day (2018) examined the flipped classroom in relation to students’ previous achievement and its effect on long-term retention and knowledge transfer in a gross anatomy class in a Doctor of Physical Therapy program. As with Morton and Colbert-Getz, Day categorized questions via Bloom’s taxonomy and found that flipped classroom students performed better on higher-level analytical questions. In addition, Day studied long-term retention of the material by analyzing student performance in the subsequent semester’s kinesiology course. Again, the flipped anatomy students performed at a higher level in the subsequent kinesiology class. Most notably, these improvements were seen to a greater extent in previously lower performing students (Day, 2018).

2.4.4.4 Assessing the Classroom Learning Environment of Flipped Classrooms

While flipped classrooms have been a popular approach within education and there has been quite a bit of research on the effectiveness of the flipped classroom, most of the research has been centered on student outcomes, specifically achievement (Bishop & Verleger, 2013; McLean, Attardi, Faden, & Goldszmidt, 2016; Rotellar & Cain, 2016). While important, student outcomes do not paint the entire picture. Flipped classrooms completely restructure the learning process, and in doing so, restructure the classroom environment. Only a few previous studies have involved the effectiveness of the flipped classroom approach using classroom learning environment criteria at the tertiary level. These will be reviewed in detail in Section 2.5.4.2.

2.5 Learning Environment and Student Attitudes

As mentioned in Section 1.2.4, Fraser (2012) defines classroom environment "in terms of the shared perceptions of the students and sometimes the teachers in that environment" (p. 1). Or, more specifically, the term learning environment "refers to the social, physical, psychological, and pedagogical context in which learning occurs and which affects student achievement and attitudes" (Afari, Aldridge, & Fraser, 2012, p. 1371). It is important to realize that learning does not happen in isolation, but is formed by the learner's overall experience, which is shaped by their learning environment (El-Sayed & El-Sayed, 2013). This section will introduce the robust field of learning environments research, which examines these shared perceptions in the classroom.

It is estimated that from primary school to the time a student completes a university degree, they will have spent over 20,000 hours in a classroom. Furthermore, abundant research on tens of thousands of students around the world has provided strong evidence that classroom environment influences student outcomes (both cognitive and affective) (Fraser, 2014, 2019). In fact, Lisso, Wilson and Simons (2002) found that at the university level, students' perceptions of their learning environment was a stronger predictor of learning outcomes than prior achievement in school. Hence, being able to assess their perceptions of their experiences in school is essential for anyone wanting to improve the learning outcomes of undergraduate classes.

In order to examine the field of learning environment research, a brief history of the field will be reviewed in Section 2.5.1, while Section 2.5.2 will examine the common goals of learning environment research, the most common methods for assessing classroom learning environments and the rationale for utilizing such techniques. Section 2.5.3 will review some historically significant instruments as well as some of the most commonly utilized in the field. Special emphasis will be given to the College and University Classroom Environment Inventory (CUCEI) which is utilized in this study (Section 2.5.4, Appendix B). Lastly, Section 2.5.5 will discuss the related field of assessing the student attitudes toward science. To assess this affective outcome, the Test of Science-Related Attitudes (TOSRA) will be discussed, which is also used in this study.

2.5.1 History of Learning Environment Research

In some respects, learning environments research is founded on earlier work that examined the interaction between an individual and their environment. Specifically, Lewin's seminal work (1936) on field theory examined an individual's interaction and their environment ("total field") and its effect on their behavior. Lewin postulated that an individual's behavior (B) is determined by the interactions of the personal characteristics of the individual (P) and that environment in which they were interacting (E). This leading to his famous formula, B=*f*(P,E), which focused future research on the need to consider behavior as a function of the individual and their environment (Fraser, 2012b, 2019).

Lewin's work was followed by Murray (1938), who proposed a needs-press model. In this model, personal *needs* are the personality characteristics of an individual that tends to motivate them to move in the direction of particular goals. Environmental *press* represents external situations that either support or frustrate the personality needs. Murray would later introduce terms to describe how to assess the environment (press). The term alpha press refers to the environment being assessed by a detached observer while beta press refers to the assessment of the environment by the perceptions of those in the environment (Fraser, 2012; 2012a; Murray & McAdams, 1938). This would later be expanded in 1956 by Stern et al., by differentiating the view of an individual within an environment (private beta press) and those views shared by members of the group (consensual beta press) (Fraser, 1998; 2012; Mishler, Stern, Stein, & Bloom, 1957).

While the work of Lewin, Murray and others set the groundwork for examining how individuals interact with their environment, the origins of learning environment research are usually attributed to the independent, seminal works of Herbert Walberg and Rudolf Moos in 1968. For his part, Walberg developed the widely used Learning Environment Inventory (LEI, see Section 2.5.3.1) out of his innovative use of classroom assessments in evaluating Harvard Project Physics (Walberg & Anderson, 1968). Walberg would go on to modify and expand Lewin's formula, B=*f*(P,E), specific to the learning. Walberg's model, L= *f*(I,A,E), has learning a function of instructional, aptitudinal and environmental characteristics (Fraser, 2019). For his part, Moos's development of his world-renowned social climate scales, involving perceptual measures in environments such as psychiatric hospitals, prisons, university residences, etc., led to the creation of the Classroom Environment Scale (CES, see section 2.5.3.2) (Fraser, 1998; Moos, 1968, 1980).

The work of Walberg and Moos on assessing the perceptions of classroom environment spawned a growth of new research in the field, as reflected by numerous articles, books, and the development of the journal, *Learning Environments Research*. The field had grown to the point that the American Educational Research Association (AERA) established the Learning Environments (LE) Special Interest Group (SIG) in 1986 (Fraser, 2014, 2019).

The early work of these visionaries can still be felt in the research done today. Moos (1976) proposed that any instrument used for assessing human environments should cover three dimensions. The first dimension relates to the interpersonal relationships of individuals in the environment and assesses how involved and supportive the members are to each other. This is known as the Relationship Dimension. The next of Moos's dimensions assesses the individual within the environment and their tendencies toward personal growth and self-improvement. This is known as the Personal Development Dimension. The third of Moos's dimensions is System Maintenance and System Change Dimension, which measures the extent to which the environment is structured in an orderly fashion and responsive to change (Fraser, 1998, 2019; Moos, 1976). The scales of the classroom environment instruments listed below (Section 2.5.3) are commonly classified according to Moos's three dimensions.

2.5.2 Assessing Classroom Learning Environments

Fraser describes the role of learning environment research as providing, "one approach for conceptualizing, assessing, investigating, and improving what goes on in the classroom" (Fraser, 2014, p. 104). A variety of analysis techniques have been utilized to assess classroom learning environments. The choice of which form of analysis to use depends on the objectives of the study, resources available, and the context in which it takes place. For example, one may choose to analyze the environment using a trained external observer (i.e., Murray's classification of alpha press, as notated above) or, one may use the perceptions of those within the environment (beta press). Again, this was further expanded upon by Stern et al. who differentiated between the perceptions of an individual within the environment (private beta press) and members of a group in the environment (consensual beta press). Furthermore, Rosenshine (1970) differentiated low inference and high inference observations. Those observations that measures very specific (often quantifiable) phenomena (e.g., how many questions a student may ask), are known as low inference, while observations that require the respondent to assess the meaning of an event in the classroom (e.g., how friendly is the teacher) are known as high inference (Fraser, 2012; Limbacher & Rosenshine, 1972).

Much of the growth of the field of learning environment research can be attributed to a specific type of analysis. It is thanks to the diverse number of valid, economical and widely-applicable assessment instruments. Fraser and Walberg (1980) outlined the following advantages for using these student-centered perceptual measures (beta press) over that of a trained observer (alpha press) (Fraser, 2012; Fraser et al., 1986):

- Economical – pencil and paper instruments are more economical for a study than hiring trained outside observers
- Based on a long period of time – instruments are based on the perceptions of students and/or teachers over many lessons rather than a few lessons that a trained observer typically views
- Pooled opinions – the results from perceptual measures include pooled opinions off all the students in the class, while observation techniques are typically view of a single person
- Student perceptions are more important than outside observer – student perceptions are determinants of their behavior (as per Lewin's B=*f*(P,E))
- Student perceptions account more for outcomes – perceptual measures have been found to account more for the variance in student outcomes than interaction variables

The growth of the field can also be attributed to the utility of these assessment instruments, and the field in general, in addressing many different research goals/objectives. For example, some of the most robust research in the field for the last 40 years has given strong evidence to the association between learning environments and student (cognitive and affective) outcomes. Not only did this research support the importance of classroom environment to improving student outcomes, but it also suggested that modifying classroom environments to better align with student preferences could lead to further improvements in student outcomes. Other common applications of these instruments include, but are not limited to: evaluating educational/curricular innovations; examining school environment as distinct from that of the individual classroom; comparing perceptions of actual versus preferred environments; evaluating if students perform better in their preferred environment (i.e., person-environment fit studies); examining differences between teacher and student perceptions; teachers' use of learning environment perceptions in guiding improvements in classrooms; evaluation of interventions to improve classrooms; typologies of classroom environments; and determining the elements of classroom environments (Fraser, 1998, 2012a, 2012b, 2014).

While there are some differences in the development and validation of the common learning environment instruments, the same overarching strategy is utilized for the majority of surveys, including the instruments detailed below in Section 2.5.3. It is a three-step process consisting of 1) identification of salient dimensions, 2) item writing, and 3) field testing and item analysis (Fraser, 2012). This basic strategy for developing personality inventory scales has been termed intuitive-rational due to the fact that the process initially identifies and defines the dimensions being assessed based on the investigators' intuitive understanding of the dimensions (Hase & Goldberg, 1967). In addition to the common form of development for many learning environment instruments, there are other common characteristics shared between many of them. Most (including those described below) are paper and pencil forms utilizing Likert ratings. Some instruments have separate forms for actual and preferred environments. Likewise, many have separate but parallel teacher and student versions. Lastly, all those listed below have scales that are categorized under Moos's 3 domains (Fraser, 1998).

2.5.3 Historically Significant and Common Learning Environment Instruments

This section will briefly review some historically significant learning and common environment instruments. These instruments were selected as they are some of the most commonly used in learning environment research. Table 2-2 lists the name of each instrument, the level for which it is intended (e.g., elementary, secondary, higher education), the number of items per scale, and each scale in the instrument categorized by Moos's three schemes. Because the College and University Classroom Environment Inventory (CUCEI) was used in my study, it is discussed in greater detail.

Table 2-2

Selected Learning Environment Instruments Adapted from Fraser (2012)

Instrument	Level	Items per scale	Scales classified according to Moos's scheme		
			Relationship Dimensions	Personal development dimensions	System maintenance and change dimensions
Learning Environment Inventory (LEI)	Secondary	7	Cohesiveness Friction Favouritism Cliqueness Satisfaction Apathy	Speed Difficulty Competitiveness	Diversity Formality Material environment Goal direction Disorganisation Democracy
Classroom Environment Scale (CES)	Secondary	10	Involvement Affiliation Teacher support	Task orientation Competition	Order and organisation Rule clarity Teacher control Innovation
Individualised Classroom Environment Questionnaire (ICEQ)	Secondary	10	Personalisation Participation	Independence Investigation	Differentiation
My Class Inventory (MCI)	Elementary	6-9	Cohesiveness Friction Satisfaction	Difficulty Competitiveness	
College and University Classroom Environment Inventory (CUCEI)	Higher education	7	Personalisation Involvement Student cohesiveness Satisfaction	Task Orientation	Innovation Individualisation
Science Laboratory Environment Inventory (SLEI)	Upper Secondary/ Higher education	7	Student cohesiveness	Open-Endedness Integration	Rule clarity Material environment
Construcitivist Learning Environment Survey (CLES)	Secondary	7	Personal relevance Uncertainty	Critical voice Shared control	Student negotiation
What Is Happening In This Classroom (WIHIC)	Secondary	8	Student cohesiveness Teacher support Involvement	Investigation Task orientation Cooperation	Equity

2.5.3.1 *Learning Environment Inventory (LEI)*

Development and validation of the preliminary version of the LEI began in the late 1960s in conjunction with the research related to the Harvard Project Physics (a new physics course for high schools at the time) (Fraser, Anderson, & Walberg, 1982; Walberg & Anderson, 1968). It was developed out of the need to find a low-cost, high inference assessment of the program. The initial assessment was known as the Classroom Climate Questionnaire and consisted of 18 scales considered meaningful for classrooms. This was later expanded into the LEI, consisting of 15 scales with seven items per scale. Each item is a statement descriptive of typical school classes. The possible responses are in a Likert scale, with respondents expressing their degree of agreement of each statement by choosing between Strongly Disagree, Disagree, Agree and Strongly Agree. For some items, the scoring direction is reversed. A typical item in the Diversity scale is: "The class has students with many different interests" and in the Formality scale is: "The class is rather informal and few rules are imposed" (Fraser et al., 1982).

2.5.3.2 *Classroom Environment Scale (CES)*

The Classroom Environment Scale (CES) was developed by Rudolf Moos and Edison Trickett to analyze the psycho-social environment of junior high and high school classes. It "conceptualizes the environment as a dynamic social system that includes not only teacher behavior and teacher-student interaction, but student-student as well" (Moos, 1980, p. 240). The CES utilizes the conceptual framework which grew out of research from a variety of human environments, including psychiatric hospitals, prisons, university residences and work milieus (Fraser, 2012a; Moos, 1968, 1976). After four different versions through field testing and item analysis, the final published version consists of 90 True-False response format items (nine scales with 10 items per scale). Example items include: "The teacher takes a personal interest in the students" (Teacher Support scale) and "There is a clear set of rules for the students to follow" (Rule Clarity scale) (Fraser, 2012a).

2.5.3.3 *Individualized Classroom Environment Questionnaire (ICEQ)*

The Individualized Classroom Environment Questionnaire (ICEQ), was designed by A. John Rentoul and Barry Fraser to measure those specific learning environment dimensions which differentiate conventional classrooms from individualized ones involving either inquiry-based or open approaches (Rentoul & Fraser, 1979). The final version of the ICEQ consists of five scales with 10 items per scale. The potential responses are a Likert format with the options of Almost Never, Seldom, Sometimes, Often and Very Often. The scoring direction is reversed for some items. Typical items are: "Students draw conclusions from information" (Investigation) and "Students work at their own speed" (Differentiation) (Fraser, 2012b; Rentoul & Fraser, 1979).

2.5.3.4 College and University Classroom Environment Inventory (CUCEI)

By the time the College and University Classroom Environment Inventory (CUCEI) was developed in 1986, the field of learning environment research had grown rapidly in primary and secondary education (Fraser et al., 1986). But while the study of the psychosocial environment of classrooms had grown rather robust at the lower levels, little analogous work had been done at the tertiary level. The CUCEI was developed to remedy this need. However, as noted by Alansari and Rubie-Davies (2020), even with the advent of the CUCEI, there is still a relative paucity of learning environment research at the tertiary level compared to that of the primary and secondary levels. As the CUCEI was utilized in this study, it will be discussed in greater detail in Section 2.5.4.

2.5.3.5 My Class Inventory (MCI)

My Class Inventory (MCI) was developed by Fraser, Anderson and Walberg (1982) as a simplified version of the LEI for use for children ages 8-12. It was then simplified again several times, evolving into a short, 25-item version. While originally developed for primary-school students, it has also been found useful for junior high students, especially for those students which may experience reading difficulties with other instruments. The MCI differs from the LEI in four important ways. First, the MCI only contains five of the LEI's original 15 scales (Cohesiveness, Friction, Difficulty, Satisfaction, and Competitiveness). This is to minimize fatigue among young children. Second, the wording of the items was simplified to enhance readability at the primary-school level. Third, the 4-point Likert scale was replaced with a two-point (Yes-No) response format. Lastly, unlike the LEI, in order to avoid potential errors transferring to a separate answer sheet, students respond on the questionnaire itself. The final form of the MCI contains either 38 items (long form) or 25 items (short form) (Fraser, 2012b; Fraser et al., 1982).

2.5.3.6 Science Laboratory Environment Inventory (SLEI)

The Science Laboratory Environment Inventory (SLEI) is one of several instruments that have been developed to evaluate a specific educational setting. In this case the SLEI, it is that of science education laboratories at the high school or higher education level. The SLEI consists of five scales with seven items per scale. The scales are: Student Cohesiveness, Open-Endedness, Integration, Rule Clarity, and Material Environment. The response options are a Likert-format ranging from Almost Never, Seldom, Sometimes, Often, and Very Often. Typical items are, "In my laboratory sessions, I do different experiments than some of the other students" (Open-Endedness) and, "My laboratory class has clear rules to guide my activities" (Rule Clarity) (Fraser, 2012a; Fraser, Giddings, & McRobbie, 1995).

2.5.3.7 Constructivist Learning Environment Survey (CLES)

Much like the SLEI and the ICEQ, the Constructivist Learning Environment Survey was created with a particular goal in mind. The CLES was developed to help researchers and teachers determine the "degree to which a particular classroom's environment is consistent with a constructivist epistemology and to assist teachers to reflect on their epistemological assumptions and reshape their teaching practice" (Fraser, 2014). The scales are: Personal Relevance, Uncertainty of Science, Critical Voice, Shared Control, and Student Negotiation with possible responses ranging from Almost Never to Almost Always. Typical items are, "I help the teacher to decide what activities I do" (Shared Control) and, "Other students ask me to explain my ideas" (Student Negotiation) (Fraser, 2014; Taylor, Fraser, & Fisher, 1997).

2.5.3.8 What is Happening In this Class (WIHIC) Questionnaire

The What is Happening in this Class (WIHIC) questionnaire was created by combining the most salient scales from a wide range of existing learning environment questionnaires with additional scales added to address contemporary concerns such as equity and constructivism. In this way, WIHIC could cover a wide-range of pertinent constructs within a single instrument, eliminating overlapping scales and duplicated items (Fraser, 1998). In doing so, the WIHIC questionnaire has become one of the most popular classroom environment instruments in the world

(Skordi & Fraser, 2019). The final version contains seven, eight-item scales including: Student Cohesiveness, Teacher Support, Involvement, Investigation, Task Orientation, Cooperation, and Equity. Typical items include, “I make friends among students in this class” (Student Cohesiveness) and “The teacher considers my feelings” (Teacher Support) (Fraser, 1998; Skordi & Fraser, 2019). The scales from the WIHIC were further used as the basis for the development for other instruments. This includes the development of an instrument that focuses on technology and outcomes in secondary school classrooms (Technology-Rich Outcomes-Focused Learning Environment Inventory, TROFLEI) (Aldridge, Dorman, & Fraser, 2004) and an instrument designed to provide teachers with feedback based on their students' perceptions of their classroom environment (Constructivist-Oriented Learning Environment Survey, COLES) (Aldridge, Fraser, Bell, & Dorman, 2012).

2.5.4 College and University Classroom Environment Inventory (CUCEI) Utilized in this Study

The purpose of this study is to explore how implementation of a flipped classroom format impacts the attitudes, perceptions of learning environment, and learning outcomes of traditional undergraduate anatomy and physiology students. Student perceptions of their learning environment were assessed using the College and University Classroom Environment Inventory (CUCEI). As described in Section 2.5.3.4, the CUCEI was initially developed in 1986 to rectify the lack of research of learning environments within higher education. The following two sections will detail the development, structure and validation of the CUCEI as well as detail how it has been utilized in studying flipped classrooms.

2.5.4.1 Development, Structure and Validation of the CUCEI

Like the primary and secondary school instruments which preceded it, the dimensions of each scale were aligned with one of Moos's three general categories of dimensions (i.e., Relationship Dimensions, Personal Development Dimensions, System Maintenance and System Change Dimensions) as detailed in Section 2.5.1. This helped guide the initial development of the CUCEI, along with creating consistency with previous secondary school instruments by using dimensions from the existing instruments to help identify potential dimensions of study for the CUCEI. Also, the goals of ensuring salience to tertiary teachers and students and wanting to create an instrument which was economical in subject answering and research processing helped guide the development of the instrument (Fraser et al., 1986). Seven scales were devised to meet the desired goals of the instrument: Personalization, Involvement, Student Cohesiveness, Satisfaction, Task Orientation, Innovation, and Individualization (see Appendix B College and University Classroom Environment Inventory (CUCEI)

NOTE: The CUCEI was developed by Fraser, Treagust and Dennis (1986) and is discussed in Sections 2.5.3.4, 2.5.4, 3.6.2, and 4.4.1. It was used in my study and included in this thesis with permission of the authors.

Initially, the preliminary version of the CUCEI contained 12 items per scale, however after field testing a sample of 127 students in 10 classes at a tertiary institution in Perth, Western Australia, some items were removed to enhance each scale's internal consistency (i.e., how much do items in a scale measure the same dimensions – with items showing low correlations with the other items of the scale being removed) and discriminant validity (i.e., how much does a particular scale measure a unique dimension not covered by other scales in the instrument – with items removed that had a higher correlation with a scale that was not its originally assigned scale). In the end, the final version of the CUCEI contains seven items per scale, hence there are 49 items total. Each item uses a four-point, Likert-like scale with the options of Strongly Agree, Agree, Disagree and Strongly Disagree with the scoring direction reversed on approximately half the items (Fraser et al., 1986).

Table 2-3

Descriptive information for each scale in the CUECI. Adapted from Fraser et al., (1986)

Scale name	Moos category	Scale descriptions
Personalization	Relationship Dimension	Emphasis on opportunities for individual students to interact with the instructor and on concern for students' personal welfare
Involvement	Relationship Dimension	Extent to which students participate actively and attentively in class discussions and activities
Student cohesiveness	Relationship Dimension	Extent to which students know, help and are friendly towards each other
Satisfaction	Personal Development Dimension	Extent of enjoyment of classes
Task Orientation	System Maintenance and System Change Dimension	Extent to which class activities are clear and well organized
Innovation	System Maintenance and System Change Dimension	Extent to which the instructor plans new, unusual class activities, teaching techniques and assignments
Individualization	System Maintenance and System Change Dimension	Extent to which students are allowed to make decisions and are treated differentially according to ability, interest, or rate of working

Once the preliminary validation data were used to create the instrument, cross validation data were collected using a larger sample of students and instructors in both Australia and the US. Internal consistency of each CUCEI scale for both the actual and preferred forms were calculated using Cronbach's alpha coefficient. The data collected suggested that across forms (i.e., actual and preferred, student and instructors, Australian and American, individual and class mean) each of the CUCEI's scales had acceptable internal consistency (Fraser et al., 1986). Discriminant validity was also examined using the larger sample measuring the mean correlation of scale with the other six scales as a convenient index. The data showed that even though the CUCEI measures distinct but somewhat overlapping aspects of classroom environment, the conceptual distinctions among scales were important enough to retain the seven dimensions within the instrument (Fraser et al., 1986).

Finally, the instrument was shown to be able to differentiate between perceptions of students in different classrooms. This was measured by examining each scale on the student actual perceptions form and performing a one-way ANOVA, with class membership as the main effect and the individual as the unit of analysis. The results showed that each scale had statistically significant differentiation between classrooms ($p < 0.001$) (Fraser et al., 1986). The CUCEI was subsequently cross-validated in New Zealand (Logan, Crump, Eonie, & Rennie, 2006).

More recently, Hasan and Fraser (2015) used a modified Arabic version of the CUCEI to assess the classroom environment of tertiary students in the UAE. Their goal was, in part, to assess the perceived environment of classrooms which had implemented activity-based teaching strategies for those who struggled with mathematics. In order to confirm the modified Arabic version of the CUCEI was valid for assessing the classroom environment of tertiary students in the UAE, a factor analysis was completed. However, in the case of the Arabic version, the scales of Student Cohesiveness and Innovation were found to be problematic during the factor analysis, therefore they were eliminated to improve reliability (Hasan & Fraser, 2015). The complete CUCEI utilized in this study can be found in Appendix B.

2.5.4.2 CUCEI and Flipped Classrooms

Strayer (2012) was the first to examine the learning environment in a flipped classroom. He utilized the CUCEI in a mixed methods study of an undergraduate statistics classroom in the United States. He found that students in a flipped statistics classroom reported to be less satisfied with the classroom structure as compared to a traditional lecture, however per the CUCEI they also were more open to the dimensions of cooperation and innovation in the classroom (Strayer, 2012). Clark et al. (2014) followed by using the CUCEI to examine how students' perceptions of their learning environment changed from before and after the flipped classroom was implemented in an undergraduate engineering class. They found a significant increase in the flipped classroom in the dimensions of individualization, innovation, involvement, and personalization (R. M. Clark, Norman, & Besterfield-Sacre, 2014).

Prashar (2015) also utilized the CUCEI to examine the learning environment in an MBA operations management course in India. There it was found that students in the flipped classroom scored higher than the traditional approach in three areas: student involvement, task orientation, and innovation (Prashar, 2015). In an engineering numerical methods course, Clark et al. (2016) used the CUCEI to compare student's perceptions of their learning environment in blended, semi-flipped and flipped class formats. They found no statistically significant differences between the groups, although they did note that the ratings for the flipped and semi-flipped modes trended lower than that of the blended (R. M. Clark, Kaw, & Besterfield-Sacre, 2016).

In one of the more recent studies, and perhaps most pertinent to this research as it examined the learning environment in a flipped classrooms of health science majors, McNally et al. utilized a revised version of the CUCEI to examine students' perceptions of their environment. Only four dimensions were included: involvement, task orientation, innovation, and cooperation. Their research differentiated students into two clusters, one group who embraced the flipped environment (they labelled "Flip endorsers") and another who did not endorse the pre-learning component (they labelled "Flip resisters"). They found that flipped endorses rated classroom involvement higher than the flipped resisters. They also found that flip endorsers rated their ideal courses higher in involvement, innovation, and cooperation, suggesting that those who are open to the flipped classroom have higher expectations (McNally et al., 2017). Unfortunately, none of these studies addressed the unique dynamics or demands typical of undergraduate anatomy and physiology classrooms.

2.5.5 Attitude Research in Science Education

As noted in the introduction, in addition to learning outcomes and perceptions of learning environment, this study also examines how flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' attitudes toward science. The subject of attitudes and its effect on behavior has a long history in education as well as the social and psychological sciences. However, the study of attitudes has been quite challenging, as attitudes exist in the brain and cannot be directly observed (Fraser, Aldridge, & Adolphe, 2010; Reid, 2015). This led many early behaviorists to suggest that attitudes could never be studied. Reid (Reid, 2015) points out that it was the creation of pencil and paper surveys to study attitudes by

Thurston in the 1920s and Likert in the 1930s which paved the way for future research. However, even with methods in place to assess attitudes, distinguishing attitudes from other constructs has been difficult. Attitudes are multi-dimensional and the definition of attitudes in the research has not been consistent (Fraser et al., 2010; Reid, 2015). One of the more recent and succinct definitions is from Eagly and Chaiken (1993, pp. 1–2), "Attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour."

2.5.5.1 Interest in Student Attitudes Toward Science

It has been said that school science in general must accomplish two goals. First, it must educate future generations about science. This goal requires that instruction include a broad overview of the domain of science. The second goal is that school science must develop the next generation of scientists. This requires a foundational knowledge of a specific discipline and its major concepts (Tytler & Osborne, 2012). Both goals are seen as essential for the advancement of society. However, since the 1960s there has been a decline in student attitudes toward science and science education, especially in advanced and industrialized societies. This decline trend is typically seen from the first year of elementary school onwards. Hence, the subject of student attitudes toward science has become a matter of concern for policy makers (Reid, 2015; Tytler & Osborne, 2012).

2.5.5.2 Assessment of Attitudes

Attitudes are thought to be stored in long-term memory and have three components:

- What we know (knowledge about the object – the cognitive component)
- How we feel (feeling about the object, like or dislike – the affective component)
- How we behave (tendency-towards-action – the behavior component) (Reid, 2015)

As attitudes reside in the brain, they must be measured indirectly by assessing behavior (assuming that behavior is an accurate measure of attitudes). To assess student attitudes in a science classroom, instruments must be created that differentiate between the aspects of attitudes. This includes being able to differentiate between "attitudes toward science" (e.g., favorable attitudes towards scientists, enjoyment of science learning experiences, development of interest in a career in science) and "scientific attitudes" (e.g., evidence as the basis of belief, acknowledging the importance of a rationale argument, skepticism toward hypotheses and claims about the natural world) (Tytler & Osborne, 2012). These can further be broken down into attitudes towards topics and themes within a particular course/subject. Therefore the assessment of scientific attitudes is really a group of many potential attitudes and this must be taken into consideration in the creation of assessment instruments (Reid, 2015; Tytler & Osborne, 2012).

Much as in the field of learning environments, surveys are commonly developed to assess student attitudes toward science. As with any instrument, the concepts of validity and reliability are essential for proper development (these concepts are described in Section 3.9.1). However, Reid (2015) points out that much of educational research in the study of attitudes does not demonstrate proper validity and reliability in the development and implementation of instruments. In addition, Tytler and Osborne (2012) cite Blalock et al. (2008) and the tendency for researchers not to use existing instruments, but to create their own instruments without subjecting them to proper psychometric development methods. Kind et al. (2007) suggest the following guidelines when creating an attitude instrument:

- Clear descriptions need to be put forward for the constructs that one wishes to measure.
- Care needs to be taken when separate constructs are combined to form one scale, with justification that these constructs are closely related.
- Reliability of the measure needs to be demonstrated by confirming the internal consistency of the construct (e.g., by use of Cronbach alpha) and by confirming unidimensionality (e.g., by using factor analysis).
- Validity needs to be demonstrated by the use of more than one method, including the use of psychometric techniques.

This study utilized a well-established, psychometrically sound instrument for assessing student's attitudes toward science which follows each of these suggested guidelines - the Test of Science-Related Attitudes (TOSRA).

2.5.5.3 Development, Structure and Validation of the TOSRA

While similar in structure to classroom environment instruments examined previously, the TOSRA was created as one component of a battery of attitude scales to assess affective outcomes in science education. Much as learning environment surveys are often associated with Moos's Social Climate Scales, the TOSRA scales are associated with Klopfer's classification scheme for affective aims in science education (Fraser, 1977, 1978). Klopfer's classification scheme, as cited in Fraser (1978), consists of six distinct categories of affective aims notated as H.1 – H.6:

- H.1 Manifestation of favorable attitudes toward science and scientists
- H.2 Acceptance of scientific inquiry as a way of thought
- H.3 Adoption of 'scientific attitudes'
- H.4 Enjoyment of science learning experiences
- H.5 Development of interests in science and science-related activities
- H.6 Development of interest in pursuing a career in science

The TOSRA builds upon a previous instrument by Fraser introduced a year prior. In both instruments, separate measures were associated with category H.1 (Manifestation of favorable attitudes toward science and scientists) as H.1 has two distinct subcategories (i.e., attitudes toward science distinct from attitudes toward scientists). Hence, seven scales were created (Fraser, 1978). The seven scales and their affiliated Klopfer's classification can be seen in Table 2-4 below.

Table 2-4

Test of Science-Related Attitudes and associated Klopfer's Classification

Scale Name	Klopfer's Classification
Social Implications of Science (S)	H.1 Manifestation of favorable attitudes toward science and scientists
Normality of Scientists (N)	H.1 Manifestation of favorable attitudes toward science and scientists
Attitude to Scientific Inquiry (I)	H.2 Acceptance of scientific inquiry as a way of thought
Adoption of Scientific Attitudes (A)	H.3 Adoption of 'scientific attitudes'
Enjoyment of Scientific Lessons (E)	H.4 Enjoyment of science learning experiences
Leisure Interest in Science (L)	H.5 Development of interests in science and science-related activities
Career Interest in Science (C)	H.6 Development of interest in pursuing a career in science

After preliminary testing and item analysis of the TOSRA, it was determined that each scale would be assessed by 10 items using a single five-point Likert response format (Strongly Agree, Agree, Disagree and Strongly Disagree) with the scoring direction reversed on approximately half the items (Fraser, 1978).

Fraser et al. (2010) point out that the TOSRA overcomes most of the commonly cited problems that it is said other attitude instruments suffer. First, each of the constructs are clearly defined. This is done by providing distinct subscales based on Klopfer's classification of students' attitudinal aims. Second, each of the scales are conceptually unique within the TOSRA, as other instruments have been found to combine conceptually different constructs to form one scale. Third, multiple past studies have given strong evidence that the psychometric quality of the TOSRA in terms of scale reliability. Lastly, past studies with large samples have demonstrated via factor analysis that each scale of the TOSRA has unidimensionality and independence (Fraser et al., 2010; Fraser & Lee, 2015).

Since its inception, the TOSRA has been shown to be very versatile. It has not only been used for assessing students' attitudes toward science, but it has also been widely adapted for other subjects. For example, Fraser and Lee (2015) point to how the TOSRA has been modified to study students' attitudes toward mathematics, chemistry, geography, Spanish and English (specifically with Chinese students learning English as a second language). In addition, Fraser and Lee (2015) point out how the TOSRA has been widely used internationally, and shown to be valid with large sample sizes in countries such as Australia, the USA, Indonesia, Singapore, Taiwan and Korea. For example, in a large, cross-national study of secondary science classrooms, a modified version of the TOSRA was used to assess 594 students from Indonesia and 567 from Australia. Only three scales were selected for the research (normality of scientists, attitude to scientific inquire, and career interest in science). Through factor analysis, the three scales were shown to have demonstrated both unidimensionality and independence (Fraser et al., 2010). In another example, 440 students were studied in Korea where attitudes toward science in different academic streams (humanities, science-oriented and science-independent) were compared. Again, a modified version of the TOSRA with only 35-items was used (all the negatively-worded items removed to reduce the time needed for administration as well as reducing any potential confusion). As with in Indonesia, a factor analysis was conducted, resulting in a 24-item 4-scale version of the TOSRA (with social implications of science, normality of scientists, attitude to scientific inquiry, and career interest in science included) (Fraser & Lee, 2015). Both examples demonstrate the versatility, usefulness and validity of the TOSRA.

As this is a well-established instrument and has been previously extensively validated – no modifications were made for this study. That said, it should be noted that the TOSRA was originally developed to analyze science-related attitudes in secondary school students. However, the TOSRA has been shown to be valid previously at the tertiary level (Ervin, 2018; Villafañe & Lewis, 2016).

2.6 Summary

This study examined how implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students. This chapter reviewed the conceptual framework for this study (Section 2.2) as well as the literature relevant to the study (Sections 2.3-2.5).

Section 2.3 examined the current state of the undergraduate anatomy and physiology classroom literature and how researchers and instructors are concerned with how to educate A&P students in the vast amount of complex information needed to prepare them for clinical work. The section reviewed many of the challenges instructors face in order to achieve this goal, including how to optimally engage with students, how to encourage the right study skills, and how to determine the best instructional strategies to ensure comprehension, overcome misconceptions, and enhance critical reasoning for clinical settings. Furthermore, instructors must determine the best resources for their students (e.g., cadavers, models, etc.) within the limitations of their institution, as well as how to make the most of the limited time available to teach the large volume of information required to be successful at the clinical level. It is especially difficult for undergraduate A&P instructors as it has been noted that the vast majority of research has been about medical school anatomy courses, and more research is needed to evaluate the curriculum of introductory A&P courses (Griff, 2016).

Section 2.4 reviewed the relatively new and promising instructional method of the flipped classroom. A brief history of the flipped classroom method was reviewed as well as some of the dominant learning theories used to explain its design. Relevant research on the method was examined, and while it has been an area of active study for the past several years, in many ways there is still a lack of consensus on exact definition of a flipped classroom as well as best standards for implementation. Research in the flipped classroom in higher education was also reviewed, as well as specifically within the health sciences, and in anatomy and physiology classrooms.

As such a new and popular instructional design, there have been numerous calls for research. Many consider much of the current research on the topic weak, and call for future research to emphasize strong experimental designs to improve the credibility and merit of the evidence for the flipped classroom approach (Evans et al., 2019). Others consider research of the flipped classroom method to be in its infancy at the university level, noting that there is a great deal of opportunity in exploring this area (Bishop & Verleger, 2013; O'Shea, 2020b). Some call for research with a stronger alignment between theoretical frameworks and the design of the curriculum, as well as research designs which allow for stronger evidence to assess higher-order thinking (Bishop & Verleger, 2013; Lundin et al., 2018; O'Flaherty & Phillips, 2015). Within the health sciences, there is a call to study the long-term impact of the flipped classroom with regard to professional practice and patient care (Chen et al., 2017). Most of the research calls for an examination of the implementation process, including what are the best strategies for determining the outside-class and in-class activities (Betihavas et al., 2016).

The next section (2.5) reviewed the rich field of learning environments and student attitudes research. The history of the field was examined as well as the strong connection between classroom learning environments and student outcomes. The various ways to assess classroom learning environments were discussed, including the common development process for popular learning environment instruments. Some of the most common learning environment instruments from an historical perspective were reviewed, with special emphasis on the learning environment survey used in this study, namely the College and University Classroom Environment Inventory (CUCEI). In addition, the related field of evaluating student attitudes toward science, again with emphasis on the instrument used in this study, the Test Of Science-Related Attitudes (TOSRA), was discussed. Within the field of learning environments, researchers are interested in how the study of learning environments and student attitudes can enlighten teachers and administrators for better student experiences and outcomes in various learning situations.

My study examined the implementation of the flipped classroom approach on an undergraduate anatomy and physiology course. As the flipped classroom method is thought to increase active learning and higher-order thinking, it may serve to solve many of the acknowledged problems facing anatomy and physiology instructors and students. However, the flipped classroom approach completely changes how the students interact with the material, the instructor and each other, and affects all three cognitive domains in the classroom. Therefore, my study includes the examination of student perceptions of their learning environment and attitudes in addition to traditional performance outcomes to achieve a fuller picture of what is happening in the classroom.

The following chapter will review the research methodology utilized in this study as well as the philosophical rationale for the methods.

Chapter 3. Methodology

3.1 Introduction

Scientific research is used to gain knowledge about the natural world, by answering questions that often are intended to find a potential relationship between two or more variables (Glass, 2014). As such, the methods of a research project must be appropriate to answer the questions and show any potential relationships between the variables in question (either correlational or causal). This chapter will review the methodological approaches used to determine if the changes in classroom dynamics, inherent in the flipped classroom method, effect student concept development, student attitudes toward the subject, and student perceptions in an undergraduate anatomy and physiology classroom by examining the learning environment and student cognitive, psychomotor and affective outcomes. In order to demonstrate credibility as well as to allow for potential replication, this chapter will detail the specific methods utilized and the philosophical rationale for these decisions.

To begin, the next section (3.2) will review the conceptual components that form the basis of the study. This includes a review of the goals of the study, the conceptual framework, the research question and how specific secondary questions were developed to create measurable constructs within the study. Next, Section 3.3 will detail the philosophical motivation leading to the choice of a mixed methods approach to address the research question.

Sections 3.4 and 3.5 will detail the research design, including the data sources and sample (Section 3.4.1), and the development of the flipped curriculum (Section 3.5.1). Section 3.6 will review the quantitative instruments utilized, while Section 3.7 will review how focus groups were developed and utilized to collect qualitative data. Section 3.8 will review how both quantitative and qualitative data were initially analyzed and show how the mixed methods technique of triangulation was used to bring enlightenment, clarification, and quality control from the quantitative and qualitative data.

Section 3.9 will outline the strategies utilized to increase the trustworthiness of the research project. As such, Section 3.9.1 will discuss the methods utilized to increase trustworthiness in quantitative data collection and interpretation. Techniques that were utilized to ensure the validity (Section 3.9.1.1) and reliability (Section 3.9.1.2) in quantitative data are reviewed. Likewise, the strategies to increase the trustworthiness of qualitative data are discussed in Section 3.9.2. This includes techniques to ensure the credibility (Section 3.9.2.1), transferability (Section 3.9.2.2) and dependability (3.9.2.3) of the qualitative data and the study as a whole.

Section 3.10 will cover the ethical considerations that needed to be addressed in order to safely and successfully conduct the study. Section 3.11 will summarize the chapter with some concluding remarks. The overall structure of this chapter is graphically depicted in the Research Design Concept Map (Figure 3-1 Research Design Concept Map). Each component of the concept map will be explained throughout the chapter.

3.2 Conceptual Components

The three conceptual components of the study, the *goals*, the *conceptual framework* and the *research questions* form a closely integrated unit (see Figure 3-2 Conceptual Components). Ultimately, goals set the foundation of the study, answering questions such as, "Why is this study worthy of doing?" and, "Why should we care about the results?" (Maxwell, 2013). The conceptual framework is based on prior research, theories, models and personal experience, and attempts to give a rationale to the potential relationships between the variables in question. Lastly, the research questions specify what is trying to be understood by the study based on the established goals and the conceptual framework. All three conceptual components are informed and framed (to some extent) by the philosophical foundations (Section 3.3).

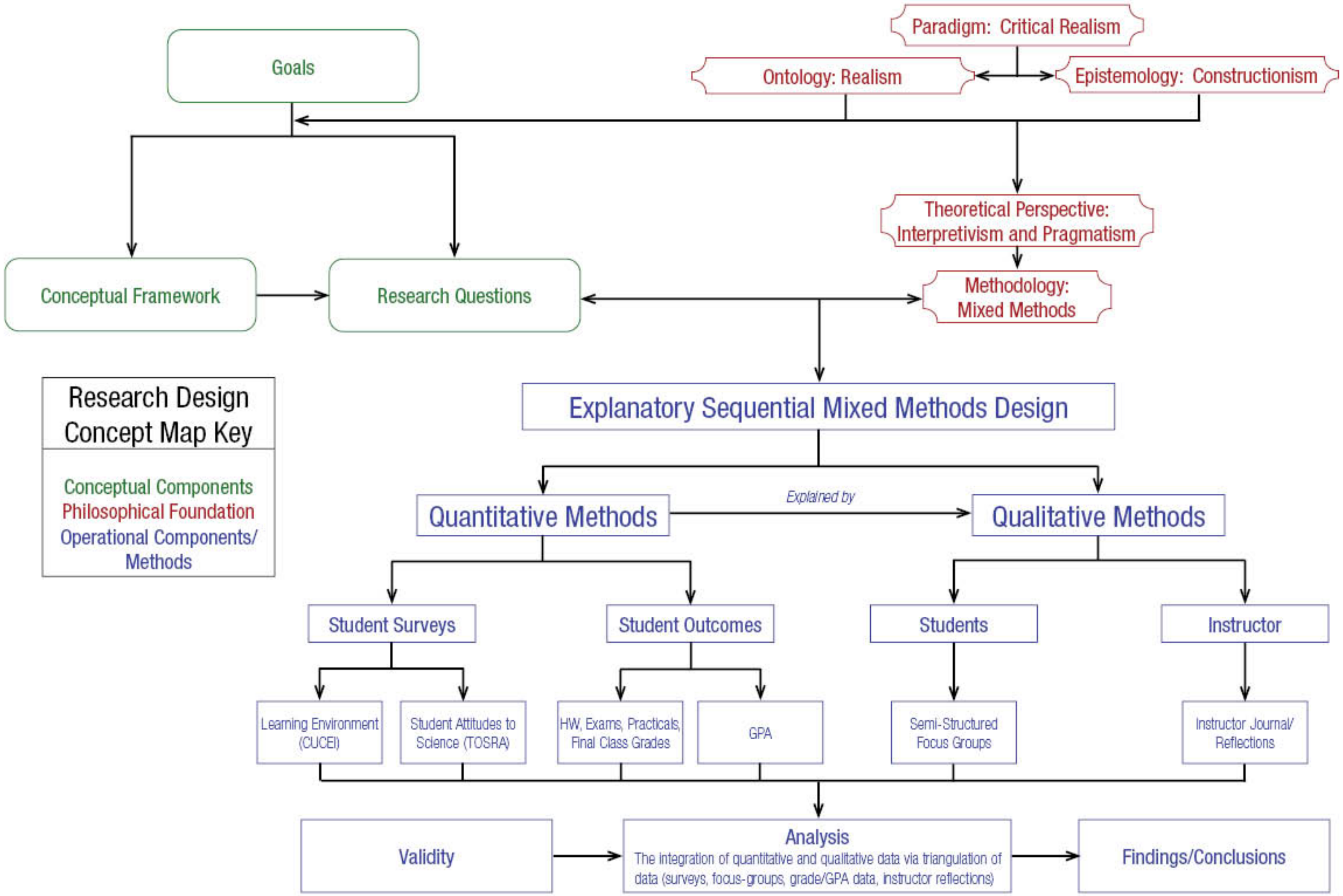


Figure 3-1 Research Design Concept Map

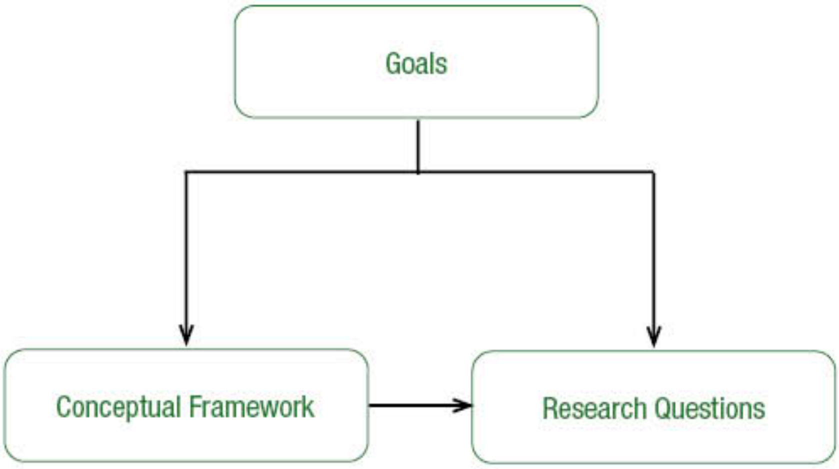


Figure 3-2 Conceptual Components

3.2.1 Goals

The unique challenges that students face in a traditional undergraduate anatomy and physiology course, combined with its importance as a seminal course in all biomedical curriculum, creates the need for innovative pedagogical approaches. As such (as discussed in Section 1.1), this research aims to determine the effect of a flipped classroom environment on the content-dense undergraduate science course: human anatomy and physiology.

Ultimately, this goal is based on my personal experience as an anatomy and physiology instructor at a private religious university as well as a creator of e-learning curriculum. This overarching goal is based on three potential outcomes:

- Improve student outcomes and overall experience in anatomy and physiology classrooms
- Better prepare students for future coursework in the biomedical studies as well as their careers in their chosen biomedical field
- Investigate potential utility of flipped classroom technique in undergraduate science courses

The goals inform the conceptual framework as well as the research questions.

3.2.2 Conceptual Framework

The conceptual framework for the study (Figure 3-3 Conceptual Framework) was introduced in Section 1.6 and detailed in Section 2.2. To review briefly, the typical challenges in undergraduate curriculum (in all three domains of learning: cognitive, psychomotor and affective) are hypothesized to be affected by the implementation of a flipped classroom format because of its unique structure. The flipped classroom format utilizes two divergent pedagogies. Traditional behaviorist principles are utilized outside the classroom in the form of video lectures, note-packets (which coincide with the lectures) and review quizzes. Constructivist principles are emphasized in class, focusing on active-learning/student-centered techniques such as group-based learning, POGILs (Process Oriented Guided Imagery Learning) and case studies. This creates a new learning environment for the students, one which changes how they interact with the information, with each other, and with the instructor. It is theorized that this new paradigm will influence each of the three learning domains and therefore addresses the stated goals.

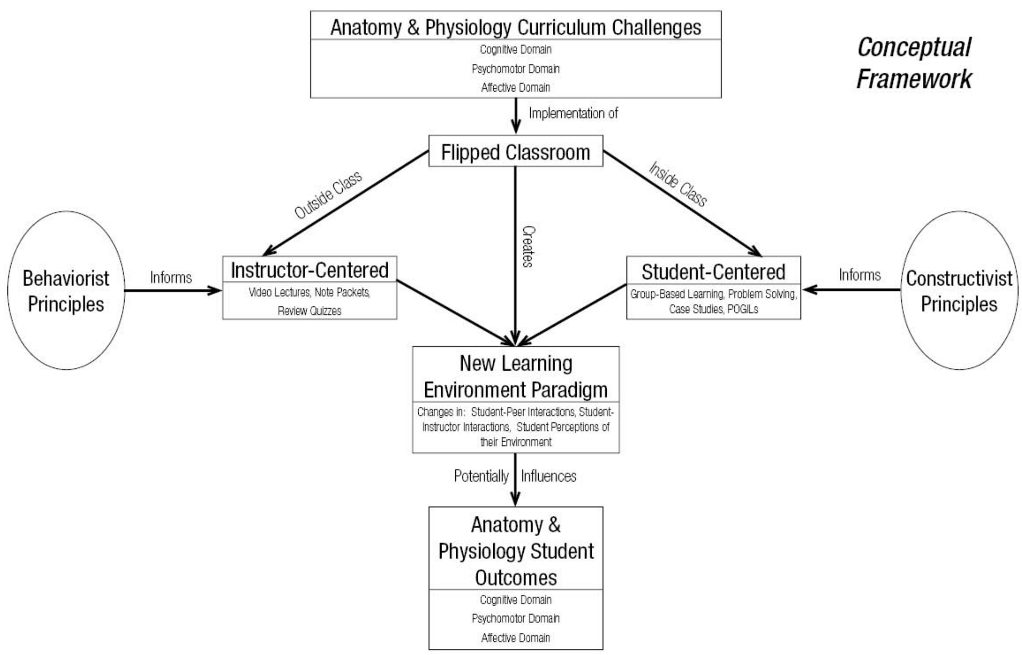


Figure 3-3 Conceptual Framework

3.2.3 Research Questions

In order develop an appropriate research question, one must address the purpose statement which takes into account the goals of the study within the conceptual framework. Taking the goals and the conceptual framework into account, the following purpose statement (as introduced in Section 1.4) was created:

The purpose of this explanatory sequential mixed-methods study is to explore how the implementation of the flipped classroom format impacted the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students.

From the purpose statement, the following primary research question was created.

What is the effect of a flipped classroom environment on undergraduate students in a human anatomy and physiology course?

In order to answer this question, and to ultimately determine the appropriate methods for data collection and interpretation (i.e., research design) we need to clarify (and specify) what we mean by the word "effect" (i.e., what "effects" are we concerned with in this research). Furthermore, this assumes that we can establish such causal relations by manipulating and measuring certain variables in the methods. Hence, the following secondary questions were established.

> *How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students':*
>
> 1) *practical laboratory performance?*
> 2) *examination achievement?*
> 3) *attitudes?*
> 4) *perceptions of their learning environment?*

Of the four components, the first two (practical laboratory performance and examination achievement) relate to the construct of student concept development. The reason for the differentiation of laboratory performance and examination achievement is that assessment of laboratory performance is primarily achieved through laboratory practicals (which emphasize memorization of anatomical structures and require a relatively greater demand on the psychomotor domain due to the activities in lab) while examinations often emphasize human physiology and require higher levels of cognitive domains. In both cases, we have clear dependent variables that operationalize the construct of student concept development, performance outcomes which, in an academic setting, are typically measured via test scores (as long as the tests have been shown to be psychometrically valid). These two components therefore focus primarily on the cognitive domain and, to a lesser extent, the psychomotor domain (with respect to laboratory skills). The next two components (student attitudes of the topic, and student perceptions of their learning environment) reflect aspects of the affective domain and are not assessed by examinations. Hence, the methods are appropriate to assess each of these domains. Put simply, these are two dependent variables that were measured through well-established learning environment and student attitude surveys which have been shown to be valid and reliable, and explained and elaborated qualitatively with focus groups.

3.3 Philosophical Foundation: Background to Methods

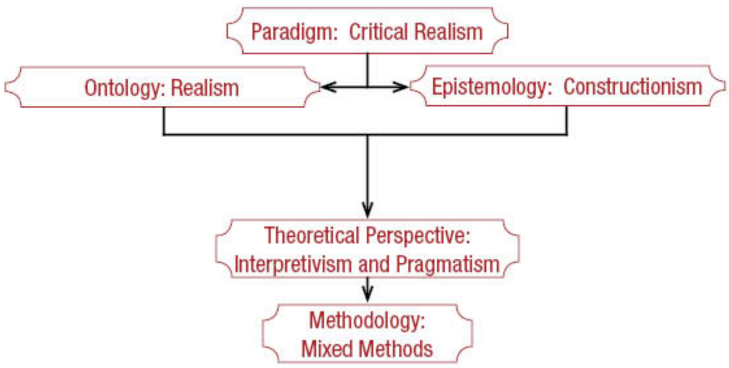


Figure 3-4 Philosophical Foundation

Scientific research is typically performed to discover something new (i.e., not currently known, often in response to some need) and/or to test someone else's (as yet unconfirmed) claim. (Glass, 2014). Specifically, as noted above, scientific research tries to answer questions and gain knowledge about the natural world (Espinoza, 2012). To understand how to properly address these questions, one must first address what we mean by "nature" and "reality" (ontology) and what we mean by "knowledge" (epistemology) (Crotty, 1998; Daniel & Harland, 2018; Glass, 2014).

3.3.1 The Importance of Establishing a Philosophical Foundation

For the research methods to be able to achieve the study's objectives and for the researcher to show that the conclusions of the study have validity, the philosophical rationale for the collection and interpretation of empirical data must be clearly articulated. Should the researcher not be clear in their attempted elucidation of a phenomenon, any hypothesized relations could be called into question and the validity of the research would be questioned.

All knowledge is built to some extent on previous knowledge with certain presuppositions that are often unacknowledged. This is also true for all research and is at the heart of every scientific question. Unfortunately, these underlying assumptions are (somewhat ironically) not often appropriately acknowledged by researchers. Authors Ben Kei Daniel and Tony Harland (2018) have shown that such philosophical underpinnings are rarely acknowledged in higher education research. Furthermore, author and qualitative researcher David Silverman seems to imply that overtheorizing such philosophical rationale can be counterproductive to conducting good research. "*I have lost count of the run of mill qualitative research papers I have come across which find it necessary to define their work in terms of obscure philosophical positions such as phenomenology or hermeneutics….. In my view, you do not need to understand these terms in order to carry out good qualitative research"* (Silverman, 2020, p. 17).

However, researchers do not work on a blank slate but rather on theories that "provide model problems and solutions" to the research community. This established framework, once again often unacknowledged, is how Thomas Kuhn defines a scientific paradigm (Kuhn, 1996). In the natural sciences, such unacknowledged presuppositions regarding the researcher's view of reality (ontology) and how it can be known (epistemology) and can unintentionally affect the research question, the method, and how data is interpreted (Glass, 2014). In the social sciences, these presuppositions exist not only in the mind of the researcher (having the same potential influences stated above) but the subjects themselves have their own presuppositions and worldviews that can potentially affect their responses in the study. As the theories being studied are relationships between phenomena, the subjects themselves are reacting and interpreting the phenomena, creating their own understanding based on their background, beliefs and values. As Maxwell (2013) states, any attempt in explaining a subject's actions without taking their perspectives into account is "probably fruitless"(p. 52). Furthermore, the researcher may adopt a paradigm regarding the worldviews (paradigm) of the subjects, hence, further influencing the interpretation of data. Unfortunately, even the most common methods for experimental design, data collection and data analysis can, "unintentionally result in inappropriate interpretations and, most problematically, in representations of data that cannot subsequently be reproduced when the experiment is repeated by others" (Glass, 2014). Hence, clarity in such matters is of the utmost importance.

3.3.2 Ontological and Epistemological Considerations

As stated above, ontology is seen as the study of being and the nature of reality (Crotty, 1998; Daniel & Harland, 2018; Espinoza, 2012). Or, as philosopher and logician Willard Van Orman Quine puts it, "It can be put into three Anglo-Saxon monosyllables: 'What is there?'" (1948, p. 21). Epistemology is the study of knowledge and how it is acquired (Crotty, 1998; Daniel & Harland, 2018; Espinoza, 2012). These two interdependent philosophical concepts set the foundation for how an individual (in our case, the researcher) understands and interprets the world. Furthermore, the ontological and epistemological position of the researcher is the basis on which their theoretical perspective and research paradigm are built. It is the theoretical perspective and the paradigm of the researcher that establish the methodologies and methods that will be used to address the research question (Crotty, 1998; Grix, 2002).

As this section should bring transparency to the research, much care must be taken to be explicitly clear in proper usage of these terms, although this can point a point of contention between various authors. For example, certain authors use the terms "objectivism" and "constructivism" as opposing ontological viewpoints (Grix, 2002) while other classify the terms as differing epistemologies (Crotty, 1998). Likewise, the terms “positivism” and “interpretivism” are defined epistemologies by some authors (Grix, 2002) while they are defined as “theoretical perspectives” by others (Crotty, 1998). Furthermore, there is often the distinction made between “philosophical ontology” (as described above) and “social ontology” which is specific to the nature of the social world. As such, I am using the following definitions for the bases of my research.

Philosophically, I am an *ontological realist*, in that reality exists, outside the human mind. However, from a social science perspective, I am an *epistemological constructivist*, in that any understanding of social phenomena will depend on the social actors involved (i.e., they “construct” meaning in social situations) (Crotty, 1998; Daniel & Harland, 2018; Grix, 2002). These may seem somewhat antithetical, however, in a sense it can be seen as somewhat analogous to Immanuel Kant’s distinction of the noumenon (the thing in itself) and the phenomena (how it appears to the observer) (Guyer, 2010) or even Plato’s *Allegory of the Cave*. While reality exists, every human constructs meaning from how they perceive their interactions with other people in a social setting. This philosophical position is often called *critical realism* (Maxwell, 2013).

3.3.3 Rationale for a Mixed Methods Approach

The metaphysical premises of the researcher discussed in the previous section determines appropriate methodologies to answer the research question. The methodologies, in turn, provide the framework for the methods. As the research question is attempting to achieve an holistic view of the potential effects of the flipped classroom (i.e., not just looking at student outcomes, or perceptions, but both and how they may be interdependent), the researcher chose a mixed methods approach that can be seen as both pragmatic (i.e., has efficacy in its practical application) and interpretivistic (i.e., taking into account the cultural and historical interpretations of social settings - see Figure 3-4 Philosophical Foundation) (Creswell, 2014; Crotty, 1998).

3.3.3.1 Definition of Mixed Methods

While Creswell (2014) acknowledges that mixed methods research can be viewed from multiple perspectives (e.g., philosophical, methodological, transformative perspectives), he defines mixed methods as a research method: "An approach to research in the social, behavioral, and health sciences in which the investigator gathers both quantitative (closed-ended) and qualitative (open-ended) data, integrates the two, and then draws interpretations based on the combined strengths of both sets of data to understand research problems" (p. 1).

3.3.3.2 A Brief History of Mixed Methods Research

The origins of mixed methods research are often seen as beginning in the late 1980s. However, the initial interest for this type of research can be seen as far back as the 1950s (Creswell, 2009; Creswell & Plano Clark, 2017). Creswell and Plano Clark (2017) define the years of early interest in mixed methods from the 1950s to the 1980s as the *formative period*. During the formative period, there were early calls for combining multiple methods (sometimes multiple quantitative methods, with others calling for triangulating both quantitative and qualitative sources of data) within the fields of psychology and sociology. According to the timeline created by Creswell and Plano Clark, the formative period overlaps with the next stage of mixed methods development, the *paradigm debate period* (1970s – 1980s). While not exclusive to mixed methods research, this was a time in the history of qualitative research where scholars debated about differing underlying assumptions between quantitative and qualitative research (Creswell & Plano Clark, 2017; Crotty, 1998). While the paradigm debate is not over, many now see the use of different paradigms as potentially complementary and can be used to address research problems (as in the case of this study, as mentioned above). Continuing the timeline, from 1980s-1990s is considered the *early procedural development period*, whereby different methods of data collection, analysis, design in mixed methods were written and solidified. Since the early 2000s, the field has become formalized, with multiple books and guidelines established for mixed methods research. This has been termed as the *expanded procedural development period* and overlaps with the *reflection and refinement period*, which, as the name implies, is a time reflecting over past controversies and concerns followed by refinements in the method (Creswell & Plano Clark, 2017).

3.3.3.3 Benefits of a Mixed Methods Approach

While not suitable for all research questions, there are times when one data source is insufficient for addressing the research question and the weakness in one source of data can be offset by the strengths of the other (Creswell & Plano Clark, 2017). This is especially true in the social sciences, whereby the problems are complex and multifaceted. Hence, the combination of techniques can give additional insight (Creswell, 2009). As described by the purpose statement and the research questions, the goal of this study is multidimensional, attempting to achieve a holistic picture of a flipped, undergraduate anatomy and physiology classroom. Hence, a mixed methods approach, utilizing the strengths of both quantitative and qualitative methods, is the most appropriate to achieve the goals of the research.

3.3.3.4 Mixed Methods in Learning Environments Research

In many ways, early forms of the mixed methods approach were utilized by learning environment researchers in the 1970s. Specifically, Kay, Trickett and Quinlan (1976) looked at teacher support and teacher control using quantitative data from students via specific scales from the Classroom Environment Scale, as well as qualitative data from outside observers (Fraser, 2012). In the Second International Handbook of Science Education (2012), educational researchers Kenneth Tobin and Barry Fraser articulate the merits of combining quantitative and qualitative data in learning environment research, with a series of successful examples dating back to 1990 (p. 1226-1228).

3.4 Operational Components: Research Design

As described above, using a mixed methods approach allows one to obtain a more comprehensive view from different perspectives. One from closed-ended responses (quantitative) and the other from open-ended personal data (qualitative) (Creswell, 2014; Fraenkel, Wallen, & Hyun, 2012). There are four important aspects to initially consider when developing mixed methods research. The first is the timing of the data collection (i.e., is it sequential, where quantitative data is collected before qualitative, or vice versa, or are they collected concurrently?). Second, one must determine the "weight" (i.e., priority) given to either quantitative or qualitative data. Third, is the concept of "mixing" which refers to when the data are merged in the process. Lastly, there is "theorizing" which refers to whether a theoretical perspective guides the research design (Creswell, 2009). These four aspects create many potential variations of mixed methods research design.

This study utilized an explanatory sequential mixed methods design. In this approach, quantitative data was initially collected and analyzed followed by a second qualitative strand that is used to explain the quantitative findings (see Figure 3-5). In doing so, the two forms of data are separate but connected, and a more complete picture is created than either method utilized in isolation. For the purposes of this study, quantitative data was collected by well-established surveys (to measure student attitudes and perceptions of their learning environment) and assessment outcomes (to measure student performance). Qualitative data was collected by semi-structured focus group interviews (to further clarify student perceptions as well as motivations) (Table 3-1) (Creswell, 2014).

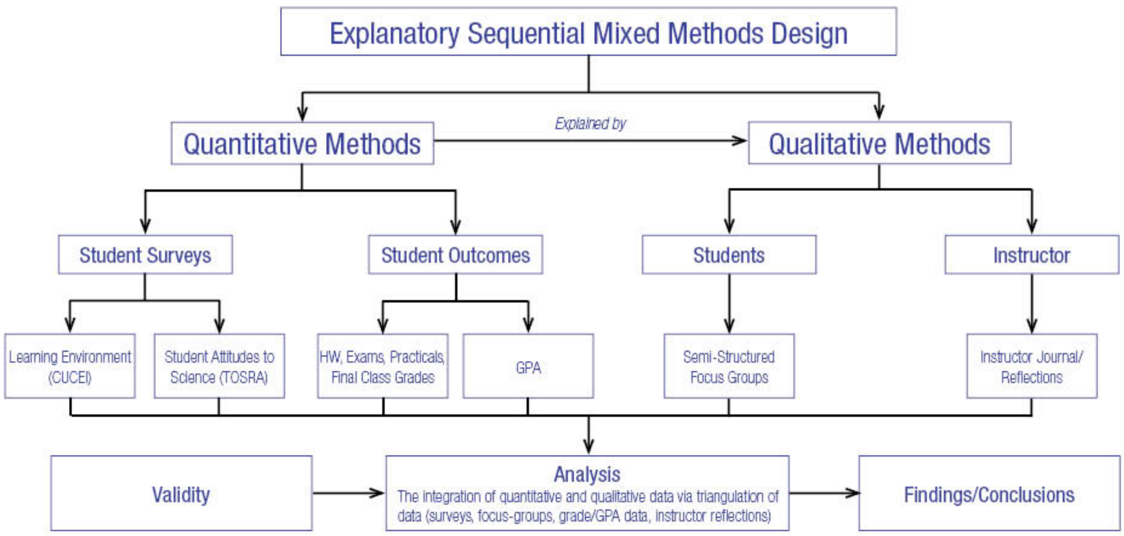


Figure 3-5 Explanatory Sequential Mixed Methods Design

Figure 3.5 positions each of the data sources for the mixed methods approach undertaken in this study. Each of these components will be discussed in detail in the sections below.

Table 3-1

Research Questions and Mixed Methods Approaches Utilized

Research Questions: How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students':	Quantitative Results	Qualitative Follow-Up Interviews
1. practical laboratory performance?	Student Performance Outcomes	Semi-structured Focus Groups, Instructor Reflections
2. examination achievement?	Student Performance Outcomes	Semi-structured Focus Groups, Instructor Reflections
3. attitudes?	Student Attitude to Science Surveys (TOSRA)	Semi-structured Focus Groups, Instructor Reflections
4. learning environment perceptions?	Student Learning Environment Surveys (CUCEI)	Semi-structured Focus Groups, Instructor Reflections

Table 3.1 connects each of the research questions to the quantitative and qualitative data used to address each.

3.4.1 Data Sources and Sample

My research took place at a small, private liberal arts university at which I am an instructor, hence the data is gathered from a convenience sample. The university at the time had an undergraduate population of 1,853 students of which 709 were male and 1,144 were female.

The university offers a two-semester course in human anatomy and physiology with multiple sections offered each 15-week semester. The class structure for each semester consists of three 50-minute lecture sections (meeting Monday, Wednesday and Friday) and one 1-hour and 50-minute laboratory section. The study is considered to be a quasi-experimental design as the students self-assign themselves into each section (typically based on what time works best in their schedule).

The study took place in the second semester of the two-semester course. The first semester was taught using a traditional lecture format. The study utilized two sections, "section-1" being a traditional lecture format (the control) while "section-3" used a flipped classroom format (the treatment). Both sections consisted of 13 undergraduate students. The sample size was smaller than expected as these were unusually small sections, as typically each section has 25 students. Section-1 consisted of 9 females and 4 males, while section-3 consisted of 8 females and 5 males. Twelve of 13 students agreed to participate from both sections. Section 3 had the unique trait of having the laboratory immediately following the lecture in the same location, hence being able to be taught as a continual 2 hour and 50-minute section for one of its three meeting times. Because of the smaller than expected sample size, data collection was considered for the following year, however due to the COVID-19 pandemic, this was not possible.

3.5 Research Design: Development of the Flipped Curriculum

The following sections will detail the development of the flipped classroom curriculum for both outside the classroom (the behaviorist component) and inside the classroom (the constructivist component).

3.5.1 Development of the Flipped Curriculum – Behaviorist Component

The flipped anatomy and physiology curriculum was developed in two overlapping stages. Initially, development began on a student note packet which served as course notes for both courses. This note packet ultimately consisted of 29 lessons over 191 pages, with 2-3 lessons typically delivered per week (see Appendix M for a sample note packet). The note packets contained many blank areas per page representing key concepts/terms for the students to fill in during the lecture. The lessons reviewed all the relevant organ systems for the semester (endocrine system, cardiovascular system, lymphatic system, immune system, respiratory system, urinary system, digestive system, and reproductive system). Several anatomy and physiology textbooks were used to develop the content of the note packets and the anatomical images within the packets were either royalty-free (via a Creative Commons license) or created by the researcher. The note packets were developed in Adobe InDesign™ and Photoshop™ and exported in a PDF form for printing and filming.

Once the initial note packets were created, filming of video lectures could proceed. Basically, the PDF note packets were opened in a computer tablet that allowed use of a digital pen. An external tablet, XP-PEN™, as well as a Microsoft Surface™ and an Apple iPad Pro™ were all utilized at various stages of development. Two recording devices were used simultaneously. I was placed in front of a green screen and was filmed as I lectured to a video camera. At the same time, screen capture software recorded the screen as I made notations as I lectured. Microphones on the camera, the computer and often a third on another device, captured the audio. After filming took place, the two video sources and the audio were synced, and the green background was removed (“keyed out”) using video editing software (TechSmith Camtasia™ or Adobe Premiere Pro™). In addition, key words were often added and animated into the video. The final product allowed the researcher to seem as if he was embedded into the note packet and the students could follow along as he made notations (Figure 3-6).

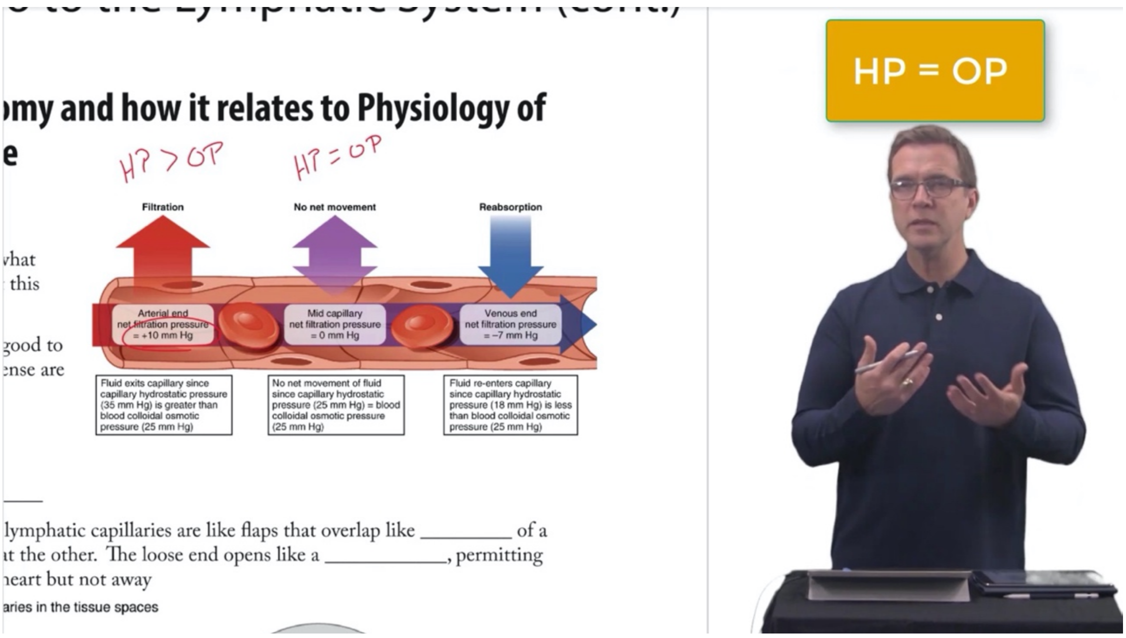


Figure 3-6 Still image of produced lecture videos

Once produced, the videos were uploaded into the university YouTube channel. The majority of the 29 note packets were presented with more than one video. As such, 84 videos were produced, ranging from 5 minutes to 28 minutes in length. Once uploaded, the YouTube URLs for each video were linked in EdPuzzle™ (https://edpuzzle.com/). EdPuzzle™ is an online platform that allows instructors to assign video lessons to their students. Video playback can be controlled (for the purposes of this study, fast-forwarding was disallowed) and questions can be embedded into the video. When the student comes to a particular place in the video at which the instructor has embedded a question, the video pauses and a question pops up to the right of the video player. All student responses are recorded in the password-secure instructor gradebook on EdPuzzle™. EdPuzzle™ embedded links were inserted into the course on the university learning management system (Blackboard™) in their respective weeks of instruction.

The note packets were made available to the students in advance in hardcopy form (picked up in class prior to online lecture) or electronically in Blackboard™. In this way, students in the flipped classroom could take notes in the note packet along with the instructor on screen. Note packets were also given to the traditional class, where they utilized them during class time lecture. The reason for note packets were twofold. First, I have found them advantageous in the past in anatomy and physiology as they include images which are often hard for students to draw (especially at the rate we need to cover them). Second, unlike PowerPoint™ presentations, as I review the note packets, the student must take handwritten notes and fill in the blanks (as opposed to taking notes on a computer). This is inspired by Mueller and Oppenheimer's (2014) classic work on the advantages of longhand note taking over laptop note taking.

3.5.2 Development of the Flipped Curriculum – Constructivist Component

In-class activities for the flipped classroom group consisted primarily of group and active learning activities, primarily consisting of POGILs (Process Oriented Guided Imagery Learning) and anatomy and physiology case studies (both are detailed below). The traditional class had note-packet lectures during class time and were given POGILs and case studies as homework. While there was still an active learning aspect to the homework, the traditional class did not have the benefit of working with peers or having the instructor for assistance while completing the assignments.

3.5.2.1 POGILs (Process Oriented Guided Imagery Learning)

POGILs were originally developed in high school chemistry classes in the mid-1990s. Rather than focusing on teacher-centric instruction with the teacher "telling" the students what they must know, POGILs were based on a three-stage learning cycle. First, POGILs are designed to take into account students' prior knowledge, including (mis)conceptions on the topic. This is done through the introduction of a conceptual model. Next, guided inquiry (i.e., carefully crafted questions) is intended to lead the student to concept invention. Lastly, students are asked to apply their new understanding through a series of problems (P. J. P. Brown, 2010). POGILs are intended to be done in groups of 6, with a formal role for each participant (e.g., a "Manager", "Spokesperson", "Scribe", "Librarian", "Reflector", and a "Consensus Builder") (M. Jensen, 2014). However, as I have been unable to successfully instruct in this format, the groups were allowed to work together in a less formal format.

3.5.2.2 Case Studies

The use of case studies is very common in health science education as they help prepare students for clinical work. Like POGILs, case studies are considered a guided inquiry approach to learning. Students are given a clinical problem and learn through group discussion. Advocates for using case studies suggest that the method encourages student exploration, allows for the correction of misconceptions, and helps focus the student's attention on key points needed for clinical work (Eseonu et al., 2013). It is also suggested that a directed case study instructional method improves the ease and depth of learning and increases appreciation of the relevance of anatomy and physiology (Cliff & Wright, 1996). The case studies utilized in the study were furnished by the publisher of the textbook (McGraw Hill™) as well as a few other examples by other publishers found by the instructor online.

3.6 Research Design: Quantitative Methods

The quantitative methods of this study consisted of student academic performance outcomes and data collected from learning environment and student attitude surveys.

3.6.1 Student Performance Outcomes

Student performance outcomes (i.e., laboratory practicals, examinations, assignment grades, etc.) are included to assess potential changes in the cognitive domain and to a lesser extent, the psychomotor domain (primarily as related the skill required to do well in laboratory practicals).

3.6.1.1 Anatomy and Physiology Examinations

Five examinations and a final were given. All five examinations were equally weighted in the gradebook and the final was worth 2 examination grades. The examinations were as follows:

- Examination 1 (Week 2): The Endocrine System
- Examination 2 (Week 5): The Digestive System*
- Examination 3 (Week 8): Blood and the Heart
- Examination 4 (Week 10): Vessels, Lymph, and the Immune System
- Examination 5 (Week 13): The Respiratory System, The Urinary System and Nutrition
- Final Examination (Week 16): 50% Reproductive System and 50% cumulative material

*NOTE: Typically, the digestive system would come later in the semester (with the other systems that exchange with the external environment), however for ease of dissection within the lab, it was decided to move the subject earlier in the semester.

The examinations were given online, utilizing the publisher's online testing platform (McGraw Hill Connect™). This allowed for a varied question-type, including but not limited to, classification questions, labeling, ranking, sequence, matching, fill-in-the-blank, multiple choice and check all that apply.

3.6.1.2 Anatomy and Physiology Laboratory Practicals

Three laboratory practicals were administered (weeks 8, 12 and 15). Each laboratory practical was weighted the same as an examination. The laboratory was divided into 24 stations. At each station there was a specimen with 2-to-4 questions. Students were assigned an initial station and had 2 minutes to answer the questions for each station. Once the two minutes were complete, they were told to rotate to the next station.

3.6.1.3 Student GPA

The original research design only called for the above examinations and practicals to be assessed, and with the addition to the learning environment surveys (discussed below in Section 3.6.2) to be clarified and expanded with qualitative data. However, upon inspection, reflection and integration of the initial quantitative and qualitative data, further quantitative data were examined to attempt to bring clarity. This was in the form of student grade point averages (GPAs). Student GPA data prior to the semester were analyzed to see if there were any correlations between a students' perception of the flipped classroom and their performance in undergraduate studies prior to class.

3.6.2 Learning Environment Data – College and University Environment Inventory (CUCEI)

Learning environment data were collected by implementation of the CUCEI (College and University Classroom Environment Inventory) survey. This, as well as student attitude toward science data (Section 3.6.3) were collected to assess components of the affective domain.

As discussed in Section 2.5.3.4, the College and University Classroom Environment Inventory (CUCEI) was originally developed and presented in 1986 to assess the classroom psychosocial environment at the tertiary level much as other instruments (e.g., LEI, CES, ICEQ, MCI) had done for primary and secondary levels successfully for years prior (Fraser et al., 1986). The CUCEI was specifically designed for use in small classes (up to 30 students) sometimes referred to as 'seminars'. Much like its predecessors, the CUCEI has four distinct forms which measure:

- Student perceptions of actual classroom environment*
- Student perceptions of preferred classroom environment
- Teacher perceptions of actual classroom environment
- Teacher perceptions of preferred classroom environment

*Only student perceptions of the actual classroom environment were examined in this study.

The survey consists of 49 items, with each item assigned to one of seven scales: Personalization, Involvement, Student Cohesiveness, Satisfaction, Task Orientation, Innovation and Individualization. Each scale is associated with one of Moos's three general categories for conceptualizing human environments with seven items per scale (see Table 2-3). Each item is responded to on a four-point scale with the options of Strongly Agree, Agree, Disagree and Strongly Disagree. For approximately half the items, the scoring direction is reversed (see Appendix B) (Fraser et al., 1986).

3.6.3 Student Attitude Toward Science Data – Test of Science-Related Attitudes (TOSRA)

As mentioned above, student attitude toward science data were collected to further assess the components of the affective domain. This was done by implementing the TOSRA (Test of Science-Related Attitudes) survey.

The TOSRA consists of 70 items. Each item is associated with one of seven scales: Social Implications of Science, Normality of Scientists, Attitude of Scientific Attitudes, Enjoyment of Science Lessons, Leisure Interest in Science, Career Interest in Science. As described in Section 2.5.5.3, the TOSRA scales are associated with Klopfer's classification scheme for affective aims in science education (Fraser, 1977, 1978). Klopfer's classification scheme consists of six distinct categories of affective aims: Manifestation of favorable attitudes toward science and scientists; Acceptance of scientific inquiry as a way of thought; Adoption of 'scientific attitudes'; Enjoyment of science learning experiences; Development of interests in science and science-related activities; Development of interest in pursuing a career in science (see Table 2-4). Potential responses on the TOSRA are on a five-point Likert format with the options of Strongly Agree, Agree, Disagree and Strongly Disagree with scoring direction is reversed on half the items (See Appendix A) (Fraser, 1978).

3.6.4 Quantitative Data Collection

Both surveys (TOSRA and CUCEI) were administered at weeks 7 and 15 of the term. This is not a traditional pre-/post- experimental design as the purpose of the two administrations was to see if the perceptions of the students in the treatment (flipped classroom) group changed over time. Unlike a pre-/post- design, this structure allowed the initial administration of surveys to measure student attitudes and perceptions after they had been exposed to the uniqueness of an anatomy and physiology classroom. Hence, the initial administration is not a traditional pretest as the subjects in the control had already been exposed to the treatment for some time.

As the researcher is also the instructor of the class, two third-party interviewers (also faculty members at the institution) administered the surveys. The reasons for this were twofold. First, this was done in order to decrease any perceived influence of the instructor on the students in regard to their participation in the study (the fact that participation in the study was completely voluntary and had no bearing in their position or grade in the class was made clear in the introduction letter which was reviewed and given to all potential subjects at this time). Second, this was done to help ensure the anonymity of the subjects. At the end of a class meeting time, the instructor/researcher left the room and the interviewers introduced themselves and the project. Each student was given an introduction letter (Appendix D) and a letter of consent (Appendix C) should they choose to participate. Each participant picked a random number out of a bag which they were instructed to put on each of their completed surveys in lieu of their name. Only the interviewers and research assistants know which number is associated with which subject (this information is kept from the researcher).

3.7 Research Design: Qualitative Methods

The qualitative methods of this study consisted of semi-structured interviews and instructor reflections.

3.7.1 Semi-Structured Focus Groups

As noted above, initial (quantitative) data collection took place during the Spring Semester (January – May, 2019). In order to allow the subjects time to reflect on their experience, semi-structured focus group interviews were conducted the following Fall Semester (after summer break). Unfortunately, this allowed for greater maturation error due to participant dropout. Specifically, a university program (Athletic Training) in which several of the students were enrolled was officially terminated during the Spring Semester. Although the university was committed in allowing all those currently enrolled to complete the program, several students transferred to other institutions over the summer and did not respond to requests to return to participate in the focus groups in the Fall.

Two focus sessions were held. Each had a mix of students from the experimental (flipped) and control (lecture) groups. The same interviewers/faculty members who originally administered the surveys, lead the interviews. A list of interview questions (Appendix D) was given to the interviewers, however they were instructed they may deviate from the questions to some extent depending on the nature of the responses. The focus groups were video recorded using Apple iPads. Each subject approached the camera with a notecard with their original number prior to the session starting. The video files were given to a research assistant to create a transcription for coding.

In order to maintain the confidentiality of the participants, the videos of each focus group session were given to a research assistant to transcribe. The transcriptions were then re-evaluated by another research assistant for accuracy.

3.7.2 Instructor Reflections

Instructor reflections were documented primarily throughout the semester of instruction in the control and the treatment classrooms. Additional reflections were documented in the following semester. The reflections were of the students' attitudes in the classroom as well as explicit, unsolicited feedback from certain students.

Instructor reflections were input into Microsoft OneNote™ on an Apple iPad™. The reflections were unstructured observations (i.e., not initially coded). Primarily, they were in the form of either inscription notes (jottings of observations as they happened) and description notes (recollections and reflections detailed after the fact). Twenty reflections were recorded.

3.8 Research Design: Analysis

This study contained both quantitative and qualitative analyses, which were integrated via triangulation.

3.8.1 Quantitative Data Analysis

All quantitative data were entered into a Microsoft Excel™ spreadsheet by a research assistant and saved as a password-protected file. A second research assistant reviewed the surveys and their entries for accuracy. After confirmation of accuracy by the second reviewer, the data was imported into R, an open-source language and environment for statistical analysis, in addition to Excel™.

As described above (Section 3.4.1) the sample consisted of two groups, a flipped and a traditional classroom. These two groups, the treatment (i.e., the flipped classroom) and the control (i.e., the traditional classroom) were compared quantitatively by analyzing classroom examinations, laboratory practicals, and GPA to assess student outcomes (addressing the research questions in the cognitive and psychomotor domains as described in Sections 1.5, 3.2.3, and 3.6.1). In addition, further quantitative data were collected to assess the students' perceptions of their learning environment and their attitudes toward science (addressing the research questions in the affective domain) by using well-established surveys (as described in Sections 1.5, 3.2.3, 3.6.2, and 3.6.3). Both surveys were given at two time intervals to look for changes in the affective domain over time.

As noted in Section 3.4.1, the sample size was much smaller than expected, with each class having only 13 students, half that of the typical enrollment. My original plan was to compare the quantitative data using common parametric techniques, such as paired and unpaired t-tests. However, this plan assumed a typical enrollment. The standard parametric techniques would not be appropriate with the small sample size and non-normal data. Therefore, non-parametric tests were utilized to analyze the quantitative data. Specifically, the Mann-Whitney U test was used to compare survey data as well as outcomes data (e.g., examinations and laboratory practicals) between the traditional and flipped classrooms. The Mann-Whitney U test is often seen as the non-parametric alternative to the independent samples t-test as they are both used to determine if there are any statistically significant differences between two unpaired groups (although it should be noted that it tests a different null hypothesis). The Wilcoxon Signed-Rank test was used to analyze the survey data of each individual group over time (weeks 7 and 15). The Wilcoxon Signed-Rank test can be seen as the non-parametric alternative to the paired Student's t-test (MacFarland & Yates, 2016a).

In addition to the non-parametric measures for statistical significance, effect sizes were also calculated. Effect sizes are an indication of the strength of the association or difference between variables (Li, 2016). Estimates of effect size are considered "useful for determining the practical or theoretical importance of an effect, the relative contribution of different factors, or the same factor in different circumstances, and the power of an analysis" (Fritz, Morris, & Richler, 2012, p. 2).

My original plan was to use one of the most common forms of estimating effect sizes between two groups known as Cohen's *d*. Cohen's *d* represents the difference between the means divided by the pooled standard deviation. Thus, the result is a number that represents the difference between the means in standard deviations. In Cohen's *d*, a calculated effect size of 0.2 is considered to be small, while an effect size of 0.5 is considered medium and an effect size of 0.8 is considered large (J. Cohen, 1988). However, Cohen's *d* requires normal data, so again due to the small sample size my original plan had to be adjusted. Instead, effect sizes were calculated using a non-parametric test which estimates the percentage of variance in the dependent variable that is explained by the independent variable. This non-parametric alternative to Cohen's *d* ($r = Z/\sqrt{n}$) (Fritz et al., 2012; Tomczak & Tomczak, 2014) was used to evaluate the differences between the two classes in learning environment scales (from the CUCEI) and student attitude scales (from the TOSRA).

As with the significance tests discussed above, effect sizes were evaluated from two different perspectives. First, to measure the overall differences in perceptions, the posttest scores were compared (i.e., the unpaired data). Second, to assess how much their perceptions changed over time, the differences in the change in scores from the initial implementation of the surveys to the posttest scores were also compared (i.e., the paired data). Other quantitative data (i.e., GPA) were evaluated after initial triangulation of data brought certain trends to the attention of the researcher.

3.8.2 Qualitative Data Analysis

Qualitative data analysis is often described as a meaning-making process (Daniel & Harland, 2018), or, put another way, "locating meaning in data" (Guest, MacQueen, & Namey, 2011 p. 49). The process involves the interpretation of data (in the case of qualitative research, data is the textual representation of a conversation, observation or interaction). This is done in a sense to simplify the data, taking the transcribed words and classifying them into categories. This is done by developing themes (i.e., a unit of meaning that is observed by the reader), and codes (a name or label which contains the boundaries or a component of a theme) (L. Cohen, Manion, & Morrison, 2018; Daniel & Harland, 2018; Guest et al., 2011).

Initial codes were derived from theoretical constructs from within the study. This is known as pre-ordinate categorization (L. Cohen et al., 2018). As suggested by Guest et al. (2011), a structural codebook was used to create an initial guide to code development. This was done by logically connecting questions of the semi-structured focus groups to learning environment scales within the CUCEI. In addition, questions were also linked to Moos's schema for assessing human environments and, when appropriate, to behaviorist or constructivist principles. The structural code includes the main question and any probes by the interviewers that were intended to enrich the responses from the subjects in the study. In this way, structural coding maps the path between theory and data collection (Guest et al., 2011). The basic structure can be seen in Figure 3-7 below.

CUCEI Scales	**Moos's Scheme**	**Behaviorist/ Constructivist**	**Question**	**Speaker ID**	**Response**
Innovation	System maintenance and change dimensions	N/A	1. Overall, how do you feel regarding the flipped classroom approach?		
Individualization	System maintenance and change dimensions	N/A	2. Compared to the first semester of Anatomy and Physiology, taught in the traditional format, which structure do you feel worked best for you?		
Personalization	Relationship Dimensions	Behaviorist	3. Do you have a preference of types of video lectures (created by the instructor or found on YouTube) used in the course?		
Involvement	Relationship Dimensions	Constructivist	4. Do you feel the activities in class assisted in your learning of Anatomy and Physiology?		
Innovation	System maintenance and change dimensions	Constructivist	5. How helpful do you feel the case studies were in learning Anatomy and Physiology?		
Student Cohesiveness	Relationship Dimensions	Constructivist	6. How helpful do you feel the POGILs were in learning Anatomy and Physiology?		
Student Cohesiveness	Relationship Dimensions	Constructivist	7. How important do you feel interacting with fellow students is to the learning process?		
Satisfaction	Relationship Dimensions	N/A	8. What would be your recommendations (if any) for the instructor to make the class better in the future?		

Figure 3-7 Initial format of the Structural Codebook

Once the transcript was incorporated into the structural codebook, initial analysis could begin. Guest, MacQueen and Namey (2011) recommend starting by reviewing the analytic objective, followed by reading the text to be analyzed and notating initial thoughts to the kinds of meaning the text may hold. This initial discovering of themes is similar to the concept of "open coding" as described by Cohen et al. (2018) as well as Daniel and Harland (2018). In open coding, one attaches a label to a piece of text, generating categories and defining properties and dimensions of each category (this is the beginning of the codebook). Ninety-four initial codes were identified by the researcher (see Appendix K). This was done by importing the structural codebook into the qualitative analysis software, Quirkos™. Each subject was created as an individual source within the software and the appropriate transcript was pasted for each subject. In addition, instructor/researcher reflections were also incorporated into the analysis software in the same manner.

Once initial themes were identified, the text was reviewed multiple times to codify themes and expand the codebook. This process took place in several stages with both the researcher and a research assistant reviewing each transcript and the assignment of codes. Results of coding between the two coders were compared, and when differences occurred, adjustments were made to the codes or the code definitions.

Initially, it was a hermeneutically oriented approach, revisiting the original text and searching for more insight into the subjects' thoughts and experiences. Much care was taken to not over-interpret the data (i.e., not "read into the data") (Guest et al., 2011). In addition, relationships between codes were acknowledged with labels created to group codes with similar meanings. This is known as axial coding. Codes were also developed to further connect the data to the phenomena being studied (i.e., flipped classrooms, learning environment perceptions, student attitudes toward science, etc.). This is known as selective coding (L. Cohen et al., 2018; Daniel & Harland, 2018). Codes were winnowed by continually reviewing the data and ensuring the codes found were relevant to the analytic objectives and truly noteworthy (i.e., those codes that seemed initially interesting to the reviewer, but ultimately just an oddity or distraction were removed). Furthermore, codes which were found to be synonymous within the context of the study were merged (Guest et al., 2011). The final list includes 74 codes (Appendix L).

The researcher moved on to formulating themes from the codes. Guest describes the entire process as clearly mapping "the path between theory and the way the data are collected, between data collection and the resulting evidence, and between the evidence and theories about what it signifies" (Guest et al., 2011, p. 75). Codes were synthesized through the conceptual framework and context of the study. Final themes are presented in Chapter 4.

3.8.3 Triangulation: The Integration of Quantitative and Qualitative Data

As noted earlier (Section 3.3.3.3), the integration of quantitative and qualitative data in mixed methods research has several advantages. It allows the researcher to not be reliant on one particular method, potentially creating bias. In addition, the more the various methods correspond, the more confidence can be held in the findings (Cohen et al., 2018). In this study, designed as an explanatory sequential mixed-methods study, the qualitiative research was originally intended to expand and explain the quantitative data. Upon evaluation of the themes derived from the qualitative data, the reseacher noticed interesting trends of student perceptions with certain types of students. This, in turn lead to the inclusion of more quantitative data (i.e., cumulative GPA prior to class) to bring clarity to the qualitative data and the study as a whole (Figure 3-8). This additional data was brought into Quarkos™ and added as properties to each subject. This allowed for separation of the emergent themes by these subject properties.

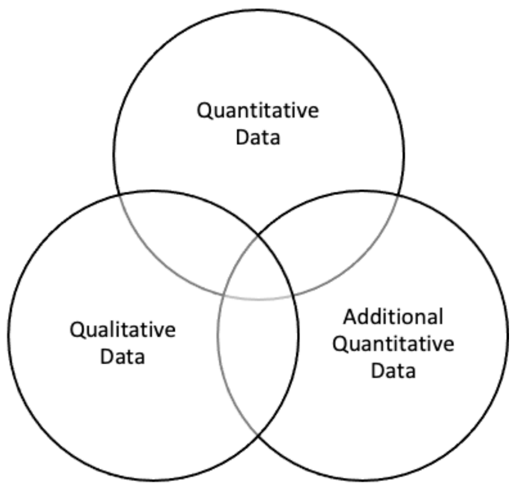


Figure 3-8 Sources of Data

3.9 Trustworthiness

This section will discuss the quality and rigor of the research. The quality of quantitative research will be discussed first, using the common criteria for evaluation of all positivistic research: validity and reliability. These criteria are not suitable for evaluating qualitative data, and as such the next section will discuss those criteria which are common to the evaluation of qualitative research (e.g., credibility, transferability, dependability).

3.9.1 Trustworthiness in Quantitative Data: Validity and Reliability

Validity and reliability are seen as the two main criteria for evaluating research across multiple disciplines (e.g., if a research project is found to be invalid, it is seen as worthless). That said, the "connotations associated with these terms are as varied as the fields that use them" (Guest et al., 2011, p. 79). Put simply, a study is considered valid when the hypothesized relation accurately reflects reality (i.e., the study is valid if the conclusions are "true"). There are several types of validity, with three of the main ones being internal validity, external validity and construct validity.

3.9.1.1 Validity

Internal validity refers to the degree one can be certain that the observed effect on the dependent variable is due to the hypothesized cause (i.e., the experimental treatment). External validity refers to the degree in which the hypothesized relation holds in general (i.e., in other groups and settings). This can be simply restated in one word: generalizability (L. Cohen et al., 2018; Guest et al., 2011). There can be several types of threat to internal and external validity. The nature of education research does not allow for the control of all possible alternative explanations, however the following were put in place to enhance the validity of the quantitative data. First, a control group was utilized, which helps mitigate the threat of maturation to internal validity (i.e., an alternate explanation formed by natural change over time). The threat of artifacts (i.e., "artificial" or unnatural reactions caused by participating in the study) can be unconsciously exhibited by both participants and the researcher. In this case, the greatest fear was that students would be concerned that their responses may affect their grades or their relationship with their instructor, who was also the researcher. To counter this, it was made clear that all responses would be anonymous and two other faculty members administered the surveys.

Construct validity is the degree in which a construct (a concept measured by a variable) is measured/manipulated accurately (i.e., did the methods actually measure and manipulate the properties that were intended?) (L. Cohen et al., 2018; Guest et al., 2011). In this project, quantitative data was in the form of student performance outcomes, learning environment surveys, and student attitude surveys (see Section 3.6). In the case of student performance outcomes (i.e., exams, laboratory practicals and classwork), the questions were developed by the researcher with 8 years experience teaching anatomy and physiology as well as from the publisher of the textbook. As for the instruments used to assess student perceptions of their learning environment as well as their attitudes toward science, both the CUCEI and TOSRA are well-established and validated surveys, having been used in multiple classrooms worldwide (see Sections 2.5.3.4 and 2.5.5.3 respectively). In addition, two research assistants were used to input data into Microsoft Excel™. One input the data and the second double-checked for accuracy.

3.9.1.2 Reliability

Reliability is related to the concept of replicability (i.e., is replication possible?). As with all education research (or social science research in general), exact replication is not possible. There are simply too many variables, (e.g., the subjects/sample, the context/setting, perceptions of reality, preciseness of measurements etc.) which cannot be exactly replicated as one would attempt in a laboratory experiment (note: laboratory research tends to maximize internal validity while field research maximizes external validity). Rather, in qualitative research the term "transferability" (described below) is used. That is not to say that in a mixed methods we cannot address the concept of the quantitative aspects of the research (e.g., how reliable were the instruments used?) (L. Cohen et al., 2018; Guest et al., 2011). For this project, reliability of the quantitative measures is related to the construct validity of the instruments utilized, which again, have historically been shown to be credible and replicable.

3.9.2 Trustworthiness in Qualitative Data: Credibility, Transferability, and Dependability

As noted above, the criteria for determining the quality of quantitative research are not deemed appropriate to judge the quality of qualitative research. In lieu of the concepts of validity and reliability, the terms most often cited for analyzing the quality of qualitative research are credibility, transferability and dependability (L. Cohen et al., 2018; Guest et al., 2011; Korstjens & Moser, 2018).

3.9.2.1 Credibility

Credibility refers to the confidence once can place in the truth of the findings (note: this is seen as the qualitative equivalent of internal validity in quantitative research). There are multiple strategies to increase credibility in qualitative research. One such strategy is prolonged engagement, which refers to lasting presence during observation or long-lasting engagement in the field with participants. The key to prolonged engagement is to allow sufficient time for the researcher to become familiar with the setting and to gather rich and sufficient data (L. Cohen et al., 2018; Korstjens & Moser, 2018). The researcher directly interacted with the subjects over 16 weeks, getting to know the students and taking research notes through the entire period.

Another strategy for increasing credibility and related to prolonged engagement is that of persistent observation. Persistent observation refers to identifying the key characteristics and relevant issues to the problem under study (and separating those observations that are irrelevant) (L. Cohen et al., 2018; Korstjens & Moser, 2018). The interactions between subjects and researcher in the classroom led to many insights which ultimately helped explain the initial quantitative data and led to the inclusion of additional quantitative data.

The use of triangulation in mixed methods research is also a strategy for increasing credibility. As Cohen et al. describe, "triangulation techniques in the social sciences attempt to map out, or explain more fully, the richness and complexity of human behavior by studying it from more than one standpoint and, in so doing, by making use of both quantitative and qualitative data" (L. Cohen et al., 2018). While there are multiple forms of triangulation, as described above (Section 3.8.3) this study used what is known as methodological triangulation, where multiple data gathering modes were used (i.e., surveys, outcomes data, focus groups, instructor reflections). In addition, investigator triangulation was also applied as several research assistants help review and analyze qualitative data (Korstjens & Moser, 2018).

3.9.2.2 Transferability

Transferability is the qualitative replacement for the quantitative concept of external validity, and hence refers to the degree to which the results of the research can be transferred to other contexts. A research project which is transferable is one where the theory generated can be understood in such a way that it can be useful in other similar situations (L. Cohen et al., 2018; Korstjens & Moser, 2018). One strategy for increasing transferability in qualitative research is that of thick description. As described by Korstjens & Moser, this description is "describing not just the behavior and experiences, but their context as well, so that the behavior and experiences become meaningful to the outsider" (Korstjens & Moser, 2018, p. 121). This research attempts to give a thick description of the methods and the context of the setting in which the data was gathered.

3.9.2.3 Dependability

Dependability is the qualitative equivalent for the quantitative concept of reliability (L. Cohen et al., 2018). This refers to the stability/consistency of the findings. In order to demonstrate dependability, one must show the particular design is in line with accepted standards. The strategy here is one of an audit trail, where transparency is demonstrated in all aspects of the project. This includes transparency in research design and implementation, in record keeping, and in development and implementation of the findings (Korstjens & Moser, 2018). This transparency is demonstrated in the detailed descriptions of all these aspects throughout this paper.

3.10 Ethical Considerations

All participants in this research were adult undergraduates enrolled in anatomy and physiology. Each student was given an introduction letter (Appendix D) explaining the purpose of the project and the procedures in detail. This was read aloud by the third-party interviewers and participants were given the opportunity to ask questions regarding the research procedures. They were then asked to read and sign an informed consent (Appendix C) which stated that their participation had no influence in their grade in the class. It was also made clear that participants were free to withdraw at any time prior to the analysis of the data.

In order to maintain confidentiality and respect the privacy of all participants in the study, all classroom environment questionnaires were anonymous and numbers were assigned randomly with all subjects (numbers were picked from a bag). Data were collected and coded by interviewers hired for the study. The researcher/instructor was not able see any identifiable raw data with the exception of grade data.

The project was approved by both the Curtin University Office of Research and Development (Appendix F) as well as Concordia University Institutional Review Board (IRB) (Appendix G).

The research generated data in the form of hard copy responses to learning environment and student attitude instruments. All other forms of data were electronic. Hard copy data are stored in a locked file cabinet and electronic data are stored on a password-protected personal computer.

3.11 Summary

This chapter detailed the methods utilized to examine how flipped classroom instruction impacted the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students. It reviewed the conceptual underpinnings of the project and discussed the philosophical rationale leading to a mixed methods approach. The chapter detailed the quasi-experimental sample utilized and the development and implementation of the flipped curriculum.

The chapter examined how, through the mixed methods approach, the study incorporated quantitative data in the form of surveys and outcomes data with qualitative data in the form of semi-structured focus groups and instructor/researcher reflections. Methods of individual analysis of the quantitative and qualitative data were described as well as the process of triangulation, which integrates the data sources in order to better understand complex phenomenon and give depth to the study.

Next, the chapter discussed methods to ensure the quality, rigor and overall trustworthiness of the study. Techniques utilized to enhance validity and reliability in the quantitative data were discussed. In addition, strategies implemented to enhance the credibility, transferability and dependability of the qualitative data and the study as a whole were reviewed. Lastly, ethical concerns were reviewed as well as the techniques utilized by the researcher to ensure that the subjects' rights were protected.

Chapter 4. Data Analysis, Results and Discussion

4.1 Introduction

This chapter will report the findings of the data collected and describe the data analysis in this study regarding how the implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of undergraduate anatomy and physiology students. After a description of the sample (Section 4.2) and a brief review of the purpose of the study, the conceptual framework and the research questions generated (Section 4.3), the findings are reviewed in five sections. The initial quantitative data collected are discussed in Section 4.4. This includes data from learning environment surveys (i.e., the College and University Classroom Environment Inventory or CUCEI), student attitudes toward sciences surveys (i.e., the Test Of Science-Related Attitudes or TOSRA), and A&P student outcomes data (e.g., examination and laboratory practical examination scores). The next section (4.5) discusses the analysis of the qualitative data gathered through semi-structured interviews and instructor reflections. Section 4.6 integrates the data from the two previous sections via the process of triangulation. Triangulation brought certain potential trends to light and more quantitative data was collected, which is reviewed in Section 4.7. This additional data was further integrated via triangulation to form a detailed picture of the experience of the students in both the traditional and flipped classrooms. Section 4.8 will describe the results as they are related to the research questions. The chapter will be summarized in Section 4.9.

4.2 Description of Sample

As described in Section 3.4.1, the study took place in the second semester of a two-semester anatomy and physiology course at a small, private liberal arts university. The first semester of the anatomy and physiology sequence was taught using a traditional lecture format. I was the instructor of the course, hence the data is gathered from a convenience sample. The study utilized two sections, “section-1” being a traditional lecture format (the control) while “section-3” used a flipped classroom format (the treatment). Both sections consisted of 13 undergraduate students. These are unusually small class sizes as typical sections of this course are full with 25 students each. Section-1 consisted of 9 females and 4 males, while section-3 consisted of 8 females and 5 males. Twelve of 13 students agreed to participate from both sections. One student in the traditional section who was participating in the project dropped out of the university in week 12, leaving only 11 participants at the end of the study for that section. A second round of data collection was considered the following year, but due to the COVID-19 pandemic, I was unable to do so.

4.3 Review of Purpose, Conceptual Framework and Research Questions

As introduced in Chapter 1, the purpose of this explanatory sequential mixed-methods study is to explore how implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students. The purpose is based on the conceptual framework (Figure 4-1), underpinning how implementation of a flipped classroom instruction model creates a new learning environment paradigm for instructors and students, potentially influencing all three domains of learning.

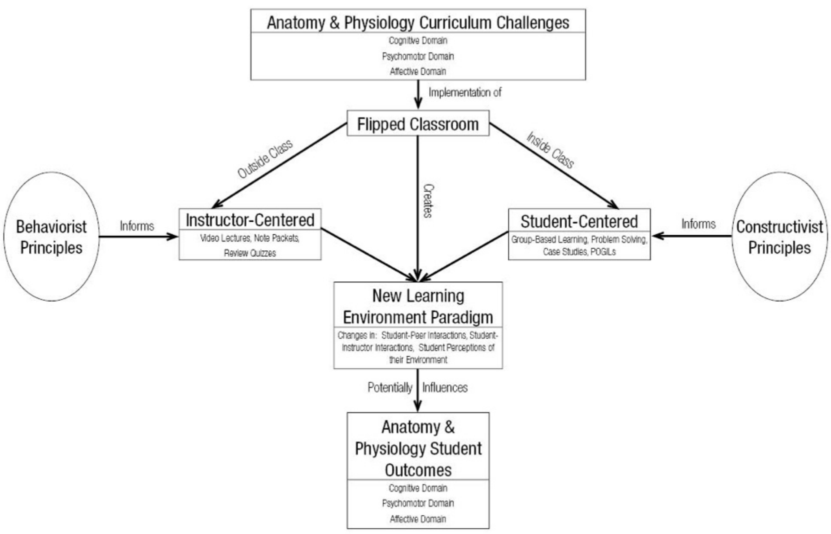


Figure 4-1 Conceptual Framework of the study

In order to achieve this goal and clearly define the investigation, the following primary research question was generated: *What is the effect of a flipped classroom environment on undergraduate students in a human anatomy and physiology course?*

From the primary research question, the following secondary research questions were generated in order to set boundaries, provide direction, and delineate measurable constructs to be investigated: *How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students'*:

1. *perceptions of their learning environment?*
2. *attitudes toward science?*
3. *examination achievement?*
4. *practical laboratory experience?*

As this was a quasi-experimental, explanatory sequential mixed-methods study, the constructs were initially quantitively measured. This was followed by qualitative data gathered through semi-structured interviews and instructor reflections. The integration of both quantitative and qualitative data was done through the process of triangulation, which brought new insights needing further clarification. This was done by collecting and analyzing further quantitative data which was again integrated and triangulated.

4.4 Quantitative Data Analysis: Comparing Instructional Groups and Changes Over Time

This section details the quantitative data gathered to examine the four constructs as articulated by the research questions (perceptions of learning environment, attitudes toward science, examination achievement, and practical laboratory experience) and described in detail in Section 3.8.1. As noted above (Section 4.2) class sizes were unusually small, and I was unable to collect any more data the following year due to the COVID-19 pandemic. Because of the small sample size, non-parametric analyses were used. Specifically, to analyze the differences between the control (the traditional classroom) and the treatment (the flipped classroom), a Mann-Whitney U test was conducted. A Wilcoxon Signed-Rank test was conducted to analyze for any differences within a single classroom over time.

Effect sizes were used in addition to significance testing to measure the magnitude of the differences between the two instructional groups as well as the differences within each group over time. As discussed in Section 3.8.1, to calculate the effectiveness of the flipped classroom for each scale of the CUCEI and TOSRA, the effect size was calculated using a non-parametric analog of Cohen's d, $r = Z/\sqrt{n}$. While Cohen's d represents the difference between the means in standard deviations, the non-parametric analog coincides with Pearson's correlation coefficient and estimates the variance in the dependent variable that is explained by the independent variable. The values of the estimated effect size range from 0 to 1.00, with below 0.3 considered a small effect, 0.30 to 0.5 a moderate effect, and above a 0.5 a large effect (Tomczak & Tomczak, 2014). As described below, there were significant differences in the scores on three scales of the CUCEI over time. Specifically, the scales of innovation (p = 0.04), involvement (p = 0.01), and individualization (p = 0.03), were higher in the flipped classroom with moderate to large effect sizes. There were no significant differences found in the TOSRA.

As also previously mentioned in Section 3.8.1, after the initial triangulation of quantitative and qualitative data brought certain trends to my attention, additional quantitative data was collected in the form of student GPA data. This additional quantitative data will be reviewed in Section 4.7.

4.4.1 College and University Classroom Environment Inventory (CUCEI) Data

In order to evaluate aspects of the affective domain of learning, the CUCEI was used to assess the students' perception of their environment in both the traditional and flipped classrooms. As discussed in Section 3.6.4, the CUCEI surveys were administered at weeks 7 and 15 of the term and then analyzed to identify potential differences between instructional method (i.e., traditional vs flipped) for the scales of the CUCEI. The average scale scores for each student can be found in Appendix E. Using the CUCEI, students' perceptions of their perceived environment were measured in seven areas:

- personalization
- involvement
- student cohesiveness
- satisfaction
- task orientation
- innovation
- individualization

R, an open-source statistical analysis language, was used to conduct both the Mann-Whitney U test and the Wilcoxon Signed-Rank test. R was also used to calculate effect sizes for differences between the treatment and the control groups as well as changes within each group over time. The results can be found in Table 4-1.

Table 4-1

Effect sizes and p-values for differences between pretest and posttest and between flipped and traditional groups for the CUCEI

Learning Environment Scale	Group	Week 7 Mean and Standard Deviation ($\bar{x} \pm s$)	Week 7 Median ($\tilde{x}$)	Week 15 Mean and Standard Deviation ($\bar{x} \pm s$)	Week 15 Median ($\tilde{x}$)	Wilcoxon Sign Rank p-value	Effect Size ($r = Z/\sqrt{n}$)
Personalisation	Flipped	4.45 ± 0.37	4.43	4.45 ± 0.49	4.50	0.88	0.14
	Traditional	4.17 ± 0.50	4.29	4.24 ± 0.36	4.14	0.67	0.18
	Mann Whitney U p-value	0.17		0.11			
	Effect Size ($r = Z/\sqrt{n}$)	0.29		0.34			
Involvement	Flipped	3.77 ± 0.62	3.50	3.65 ± 0.50	3.57	0.42	-0.25
	Traditional	3.20 ± 0.56	3.21	3.09 ± 0.49	2.93	0.64	-0.20
	Mann Whitney U p-value	0.06		0.01			
	Effect Size ($r = Z/\sqrt{n}$)	0.40		0.52			
Student Cohesiveness	Flipped	3.86 ± 0.63	3.79	3.71 ± 0.69	3.64	0.57	-0.18
	Traditional	3.42 ± 0.73	3.50	3.29 ± 0.81	3.36	0.81	-0.09
	Mann Whitney U p-value	0.26		0.35			
	Effect Size ($r = Z/\sqrt{n}$)	0.24		0.20			
Satisfaction	Flipped	4.15 ± 0.50	4.07	4.07 ± 0.54	4.07	0.50	-0.22
	Traditional	3.80 ± 0.63	3.86	3.63 ± 0.69	3.79	0.26	-0.35
	Mann Whitney U p-value	0.22		0.11			
	Effect Size ($r = Z/\sqrt{n}$)	0.26		0.34			
Task Orientation	Flipped	4.26 ± 0.51	4.21	4.06 ± 0.47	4.06	0.08	-0.53
	Traditional	4.00 ± 0.52	4.07	3.81 ± 0.56	3.93	0.33	-0.37
	Mann Whitney U p-value	0.31		0.24			
	Effect Size ($r = Z/\sqrt{n}$)	0.21		0.25			
Innovation	Flipped	3.39 ± 0.60	3.29	3.19 ± 0.57	3.14	0.33	-0.23
	Traditional	2.89 ± 0.38	2.93	2.51 ± 0.54	2.50	0.06	-0.63
	Mann Whitney U p-value	0.03		0.04			
	Effect Size ($r = Z/\sqrt{n}$)	0.46		0.43			
Indivdiualization	Flipped	2.93 ± 0.35	2.86	3.08 ± 0.35	3.08	0.14	0.39
	Traditional	2.70 ± 0.42	2.71	2.64 ± 0.42	2.43	0.67	-0.18
	Mann Whitney U p-value	0.26		0.03			
	Effect Size ($r = Z/\sqrt{n}$)	0.24		0.46			

Interpretation of effect sizes: 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

Moderate Effect Size | Large Effect Size | p-value < 0.05

4.4.1.1 CUCEI Scales Showing Significant Differences Between Instructional Groups

Once the data were collected and analyzed, three learning environment scales on the CUCEI in Table 4-1 demonstrated significant differences in the mean scores between the flipped and traditional classrooms in the final (Week 15) measurement. Boxplots were created for each of the three scales to visualize the differences in the mean student scale scores between the two classrooms. Boxplots are commonly used to graphically represent the measures of central tendency and overall dispersion of a numerical variable. The bottom line of the box represents the 25th percentile, the middle line represents the 50th percentile (the median), and the top line represents the 75 percentile. "Whiskers" are horizontal lines that represent data that go beyond the 25th and 75th percentiles. Outliers are represented by small circles beyond the whiskers (MacFarland & Yates, 2016b).

The first scale was related to involvement (see Figure 4-2). The mean student score of the learning environment scale of involvement was found to be significantly higher in the flipped classroom than in the traditional classroom with a large effect size ($p = 0.01$; $r = 0.52$). This scale is associated with the relationship dimension of Moos's scheme, and is described as the "extent to which students participate actively and attentively in class discussions and activities" (Fraser et al., 1986, p. 46). This seems logical as the in-class activities of the flipped classroom were based on constructivist principles, emphasizing active learning.

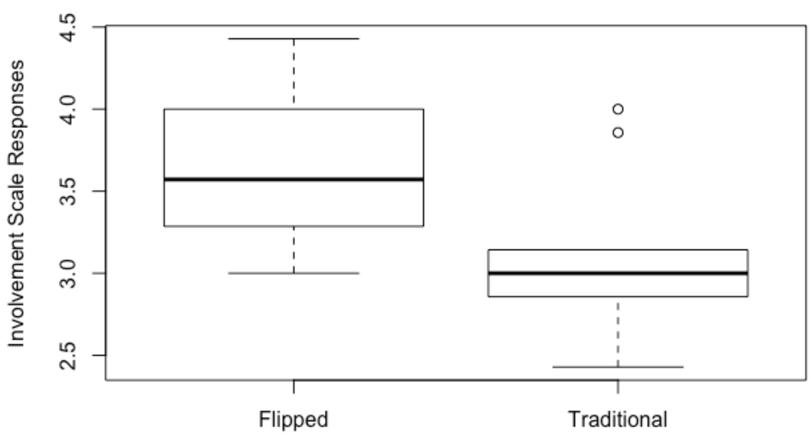


Figure 4-2 Boxplot of CUCEI Scale of Involvement for both Instructional Groups

The next CUCEI scale that showed significant differences in the mean scores between the two groups was that of innovation (see Figure 4-3). The mean student score of the learning environment scale of innovation was unique in that it was found to be significantly higher in the flipped classroom than in the traditional classroom in both the initial measurement (Week 7, p = 0.03; r = 0.46) and the final measurement (Week 15, p = 0.04; r = 0.43) with a moderate effect sizes both times. Innovation is associated with Moos's dimension of system maintenance and change. It is described as the "extent to which the instructor plans new, unusual class activities, teaching techniques and assignments" (Fraser et al., 1986, p. 46). Again, this seems logical as the flipped classroom was new to the students and used various technologies to which they were unaccustomed.

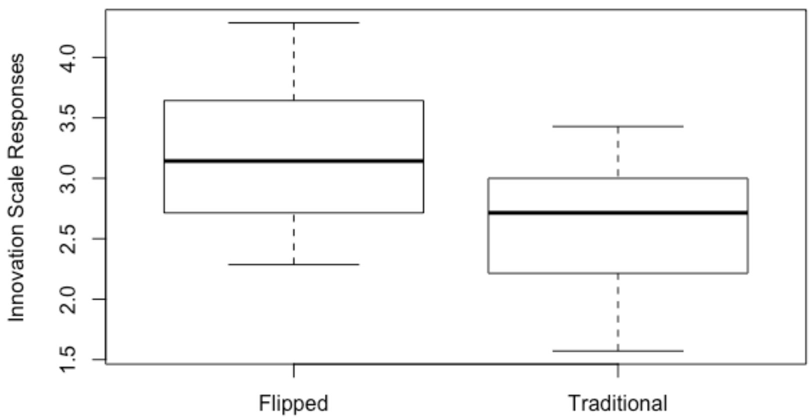


Figure 4-3 Boxplot of CUCEI Scale of Innovation for both Instructional Groups

The final CUCEI scale showing significant differences in mean scores between the treatment and the control groups was individualization (see Figure 4-4). The mean student score of the learning environment scale of individualization was found to be significantly higher in the flipped classroom than in the traditional classroom with a moderate effect size (p = 0.03; r = 0.46). Like innovation, individualization is also associated with Moos's dimension of system maintenance and change. It is described as the "extent to which students are allowed to make decisions and are treated differentially according to ability, interest or rate of working" (Fraser et al., 1986, p. 46). Again, this makes sense as the structure of the flipped classroom is intended to facilitate deep learning and help students recognize and understand their own learning styles and preferences (McLean et al., 2016).

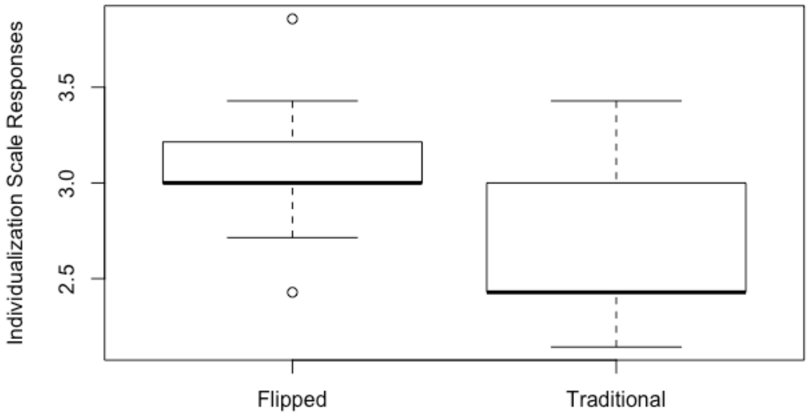


Figure 4-4 Boxplot of CUCEI Scale of Individualization for both Instructional Groups

4.4.1.2 CUCEI Scales Showing No Significant Differences Between Instructional Groups

There were four learning environment scales on CUCEI that show no significant differences in student mean scores between the flipped and traditional classrooms (see Table 4-1). These scales were personalization, student cohesiveness, satisfaction, and task orientation. The first three are associated with Moos's dimension of relationship while task orientation is associated with the dimension of personal development.

The learning environment scale of personalization is described as "emphasis on opportunities for individual students to interact with the instructor and on concern for students' personal welfare" (Fraser, 1978, p. 46). With this description, I was a little surprised that there was not a significant difference between groups, as the nature of a flipped classroom is intended to increase interactions between instructor and student (McLean et al., 2016). However, it can be noted that there was a moderate effect size for the differences between between the flipped and traditional classrooms in Week 15 ($r = 0.34$), with the flipped group scoring the scale higher than the traditional.

Likewise, the learning environment scale of satisfaction did not show a significant difference between the two classrooms but did show a moderate effect size in Week 15 ($r = 0.34$) with the flipped classroom scoring higher than the traditional. Satisfaction is defined on the CUCEI as "extent of enjoyment of classes" (Fraser, 1978, p. 46). From the qualitative data collected from instructor reflections and semi-structured focus groups, it did seem that the majority of the students in the flipped classroom enjoyed the experience. This will be discussed in Section 4.5.

The remaining learning environment scales on the CUCEI of student cohesiveness and task orientation showed no significant differences and no moderate to large effect sizes. Student cohesiveness is defined as the “extent to which students know, help and are friendly towards each other” (Fraser, 1978, p. 46). I was a little surprised that the flipped classroom did not score higher here, as students have more opportunity to interact in the flipped classroom format (Arner, 2020). Task orientation is the only scale of the CUCEI that is associated with Moos’s dimension of personal development and is defined as “Extent to which class activities are clear and well organized” (Fraser, 1978, p. 46). As I later reviewed my instructor reflections (detailed in Section 4.5.1), I realized that I had spent so much time and energy on the outside class activities that I neglected to spend enough time developing the in-class activities (see Theme G in Section 4.5.2.3).

4.4.1.3 Changes in CUCEI Scales Over Time

As the flipped classroom approach is a relatively new instructional method and was new to most of the students, I was interested to see if the perceptions of those in the flipped classroom changed throughout the semester. The data in Table 4-1 showed there were no significant differences in any of the learning environment scales within either classroom over time. For the traditional classroom, this should be expected as there were no instructional changes implemented over the semester.

4.4.1.4 Trends in CUCEI Data

While there were no significant differences in CUCEI scales over time, there are some interesting trends in effect sizes in Table 4-1. The traditional group scored the scales of satisfaction and innovation lower with a moderate and large effect size, respectively. It is unclear why this is as the only difference between groups was the nature of the topics. A possible explanation is that later in the semester are some challenging topics from a physiological standpoint, which may explain the lower perceptions. It should be noted that the flipped classroom also scored lower on these scales, just with a small effect sizes.

The flipped classroom was scored lower for task orientation over time with a large effect size. Again, nothing changed with the structure of the class. However, later in the term, I was having challenges in development of the flipped classroom curriculum (see Theme A in Section 4.5.1.1), so it was not as well structured as earlier in the semester.

The flipped classroom also was scored higher over time for individualization, with a moderate effect size. What was more interesting, however, was that while the flipped classroom was scored higher over time on individualization, the traditional classroom was scored lower. This was the only scale on the CUCEI to have one group increase over time while the other group decreased. It may be the case that, while the topics became more difficult, the flipped classroom allowed more individual attention on these tough topics.

4.4.2 Test Of Science-Related Attitudes (TOSRA) Data

In order to further evaluate the affective domain, student attitudes toward science in both the flipped and traditional classrooms were assessed by the administration of the TOSRA. As detailed in Section 3.6.4, the TOSRA surveys were administered with the CUCEI on weeks 7 and 15. The average scale scores per student can be found in Appendix F. As detailed in Section 2.5.5.3, the TOSRA measures seven scales that are each associated with Klopfer's classification scheme for affective aims in science education (Fraser, 1977, 1978). TOSRA scales include:

- Social Implications of Science (S)
- Normality of Scientists (N)
- Attitude to Scientific Inquiry (I)
- Adoption of Scientific Attitudes (A)
- Enjoyment of Scientific Lessons (E)
- Leisure Interest in Science (L)
- Career Interest in Science (C)

As with the CUCEI, the differences between the two classrooms in TOSRA scale scores were analyzed using the Mann Whitney U test using R, while differences between week 7 (initial) and week 15 (final) administrations of the TOSRA were compared using the Wilcoxon sign rank test in R. Effect sizes were also calculated in R. Table 4-2 shows that there were no significant differences on any of the seven scales between the flipped and traditional classroom in either the initial or the final administrations of the instrument. In addition, Table 4-2 shows that there were no significant changes in either group over time.

4.4.2.1 Trends in TOSRA Data

While there were no attitude scales showing significant differences between instructional methods in TOSRA responses, one scale (social implications of science) did show the flipped classroom scoring higher than the traditional classroom with a moderate effect size in Week 15 (r = 0.34). Social implications of science is based on Klopfer's classification of "manifestation of favorable attitudes toward science and scientists" (Fraser, 1978, p. 510). One potential explanation is that the flipped classroom had more interactions between students and between students and instructor (as demonstrated by the significantly higher involvement score on the CUCEI).

In addition, there were several scales in Table 4-2 for which attitude scores increased for the flipped classroom and decreased for the traditional classroom over time: normality of scientists, adoption of scientific attitudes, and career interest in science. While none of the effect sizes for changes over time were large or moderate, the differences in the effect sizes between classrooms were moderate. Again, these changes over time may be due to the increased difficulty in the curriculum later in the term with the treatment group feeling better supported due to the nature of the flipped classroom.

Table 4-2

Effect sizes and p-values for differences between pretest and posttest and between flipped and traditional groups for the TOSRA

Attitude Scale	Group	Week 7 Mean and Standard Deviation ($\bar{x} \pm s$)	Week 7 Median ($\tilde{x}$)	Week 15 Mean and Standard Deviation ($\bar{x} \pm s$)	Week 15 Median ($\tilde{x}$)	Wilcoxon Sign Rank p-value	Effect Size ($r = Z/\sqrt{n}$)
Social Implications of Science	Flipped	4.16 ± 0.42	4.15	4.22 ± 0.48	4.30	0.37	0.27
	Traditional	3.99 ± 0.37	3.90	4.02 ± 0.29	4.00	0.94	0.03
	Mann Whitney U p-value	0.26		0.11			
	Effect Size ($r = Z/\sqrt{n}$)	0.24		0.342 *			
Normality of Scientists	Flipped	3.66 ± 0.59	3.60	3.77 ± 0.56	3.80	0.36	0.22
	Traditional	3.64 ± 0.39	3.60	3.53 ± 0.58	3.30	0.61	-0.15
	Mann Whitney U p-value	0.98		0.34			
	Effect Size ($r = Z/\sqrt{n}$)	0.01		0.21			
Attitude to Scientific Inquiry	Flipped	3.46 ± 0.73	3.45	3.45 ± 0.82	3.45	0.91	-0.05
	Traditional	3.56 ± 0.69	3.50	3.59 ± 0.80	3.60	0.69	0.13
	Mann Whitney U p-value	0.76		0.71			
	Effect Size ($r = Z/\sqrt{n}$)	-0.07		-0.08			
Adoption of Scientific Attitudes	Flipped	3.93 ± 0.45	3.90	4.02 ± 0.47	4.05	0.67	0.14
	Traditional	3.90 ± 0.43	3.90	3.78 ± 0.39	3.90	0.68	-0.18
	Mann Whitney U p-value	0.95		0.39			
	Effect Size ($r = Z/\sqrt{n}$)	0.02		0.19			
Enjoyment of Science Lessons	Flipped	4.18 ± 0.74	4.25	4.18 ± 0.63	4.30	0.97	-0.06
	Traditional	4.19 ± 0.46	4.00	4.20 ± 0.54	4.00	0.72	0.18
	Mann Whitney U p-value	0.73		0.98			
	Effect Size ($r = Z/\sqrt{n}$)	-0.08		-0.01			
Leisure Interest in Science	Flipped	3.26 ± 1.05	3.30	3.46 ± 0.77	3.60	0.54	0.10
	Traditional	3.42 ± 0.46	3.30	3.48 ± 0.51	3.40	0.48	0.23
	Mann Whitney U p-value	0.85		1.00			
	Effect Size ($r = Z/\sqrt{n}$)	-0.05		0.00			
Career Interest in Science	Flipped	3.49 ± 0.86	3.70	3.59 ± 0.63	3.70	0.53	0.18
	Traditional	3.52 ± 0.69	3.50	3.44 ± 0.57	3.40	0.47	-0.26
	Mann Whitney U p-value	0.71		0.51			
	Effect Size ($r = Z/\sqrt{n}$)	-0.08		0.15			

Interpretation of effect sizes: 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

Moderate effect size *

Another trend of note in Table 4-2 was the overall absence of sizable effect sizes. While there may be several reasons for this, including the nature of such a small sample size, one other is the fact that the TOSRA measures attitudes in general rather than attitudes toward the subject of anatomy and physiology specifically. In the past, the TOSRA has been modified to measure mathematics in the form of the TOMRA (Test Of Mathematics Related Attitudes) (Spinner & Fraser, 2005). Future studies may consider modifying the TOSRA in a similar way so the items are specific to measuring students' attitudes of the subject of anatomy and physiology.

4.4.3 Learning Outcomes Data

As reflected in the research questions, initial learnings outcome data was collected in two areas, lecture examination grades and laboratory practical examination grades.

4.4.3.1 Examination Data

Examination results were analyzed to assess the cognitive domain. There were five examinations administered during the semester as well as a final examination which was twice as long in both time and number of items as each of the previous examinations and worth twice as much in the gradebook.

- Examination 1 (Week 2): The Endocrine System
- Examination 2 (Week 5): The Digestive System*
- Examination 3 (Week 8): Blood and the Heart
- Examination 4 (Week 10): Vessels, Lymph, and the Immune System
- Examination 5 (Week 13): The Respiratory System, The Urinary System and Nutrition
- Final Examination (Week 16): 50% Reproductive System and 50% cumulative material

Examination data for each student can be found in Appendix G and the means and standard deviations for each class can be found in Table 4-3. Scores on each examination were compared for instructional groups via a Mann-Whitney U test. R was also used to calculate the Mann-Whitney U test as well as effect sizes for differences between the treatment and the control groups. There was no significant difference of examination performance between the two instructional groups. The traditional class performed higher on average for every examination than the flipped class with a small effect size on all but one examination (examination 2). As the instructor, this was not surprising to me. Being a quasi-experimental design with a relatively small sample, it happened that some of the highest performing students in the cohort signed up for the traditional section as a group. This insight was important during the initial triangulation of results as explained below in Section 4.6.

Table 4-3

Effect sizes and p-values for differences between flipped and traditional groups for mean examination scores

Examination	Flipped Mean and Standard Deviation	Traditional Mean and Standard Deviation	Mann Whitney U p-value	Effect Size (r = Z/√n)
Examination 1: The Endocrine System	75.91 ± 13.38	79.06 ± 13.5	0.39	0.18
Examination 2: The Digestive System	75.65 ± 14.82	79.33 ± 11.65	0.69	0.09
Examination 3: Blood and the Heart	76.00 ± 12.60	79.57 ± 13.83	0.39	0.18
Examination 4: Vessels, Lymph, and the Immune System	79.51 ± 12.03	84.05 ± 10.29	0.34	0.20
Examination 5: The Respiratory System, The Urinary System and Nutrition	65.88 ± 10.00	72.54 ± 11.11	0.15	0.29
Final Examination: 50% Reproductive System, 50% Cumulative Material	72.50 ± 14.37	76.55 ± 13.78	0.54	0.13

Interpretation of effect sizes: 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

4.4.3.2 Laboratory Practical Examination Data

Laboratory practical examinations assess both the cognitive and psychomotor domains. The psychomotor domain is applicable here because of the laboratory skills needed to be successful on the practical examinations (e.g., dissection skills, working with physiology instruments/technology, etc.). Unfortunately, the nature of the practical examinations does not allow the disentanglement of these two domains, as the raw score does not differentiate between items that have a high psychomotor component and those that do not. The practical examination data can be found in Appendix G and the means and standard deviations for each class can be found in

Table 4-4. Traditional and flipped classrooms were compared using the Mann-Whitney U test. As with the examination data, R was also used to calculate the Mann-Whitney U test as well as effect sizes for differences between the instructional groups. There no significant difference was found between the two classrooms. As with the examinations described above (Section 4.4.3.1), the traditional class scored higher on average with a small effect size on each of the practical examinations (see Table 4-4).

Table 4-4

Effect sizes and p-values for differences between flipped and traditional groups for mean practical examination scores

Examination	Flipped Mean and Standard Deviation	Traditional Mean and Standard Deviation	Mann Whitney U p-value	Effect Size (r = Z/√n)
Practical Examination 1	43.58 ± 9.52	48.77 ± 10.13	0.30	0.21
Practical Examination 2	58.04 ± 20.06	69.21 ± 16.67	0.41	0.29
Practical Examination 3	27.5 ± 6.74	30.29 ± 5.50	0.41	0.17

Interpretation of effect sizes: 0.10 - < 0.3 (small effect), 0.30 - < 0.5 (moderate effect) and >= 0.5 (large effect)

4.5 Qualitative Data Analysis and Results

Qualitative data consisted of instructor reflections collected before, during and after the semester and semi-structured focus groups conducted the following semester. Qualitative data was imported into Quirkos™ for analysis.

4.5.1 Instructor Reflections

Instructor/researcher reflections were analyzed first. This was because the semi-structured focus groups took place the following fall semester, so my reflections could be analyzed in the summer. Reviewing my notes and reflecting on the semester, several themes emerged. As described in Section 3.8.2, the development of the themes below was a multi-step process in which the codes generated were analyzed within the context and conceptual framework of the study. Codes were merged into initial themes, which again could be collapsed into a larger theme. This process continued until the unique themes below emerged which could not be collapsed any further.

4.5.1.1 Theme A: Difficulties creating and implementing the flipped classroom

First, the challenges of implementing the flipped classroom made the semester very challenging, which is not uncommon for others who have implemented the flipped classroom approach in the past (O'Shea, 2020a). I struggled on how to best create quality videos with appropriate questions, how to properly align the curriculum outside of class with that inside the class, and how to change from my very comfortable role of lecturing ("sage on the stage") to that of "guide on the side". Often, filming and editing were completed just before needing to be uploaded to Blackboard. The process always seemed to take much longer than I'd anticipate, often besieged by unforeseen technical issues. Implementation of the flipped classroom was found to be incredibly challenging due to all the components required to implement.

4.5.1.2 Theme B: Difficulties with student participation

It was also apparent in my reflections that while certain students seemed to enjoy the flipped classroom format, others did not "buy into" the concept and were not engaged in the process. While I noticed lack of proper participation from several students, one student in particular stood out. I notated several times that this one student rarely completed the outside class video lecture assignments (EdPuzzles™) on time. He initially complained that the note packets, which were handed out as hard copies in advance in class, were not available electronically as he preferred to take notes on his iPad (which was his reason for not completing them on time). I would put the note packets in PDF form on Blackboard for him, but he would still not complete the assignments on time, if at all. This one student simply refused to participate.

4.5.1.3 Theme C: Negative preconceptions of flipped classroom

Only once did a student verbalize directly to me how much they disliked the concept of the flipped classroom model. Ironically, the person was not in the flipped classroom section, but rather the traditional. I recall how passionate she was in her view and I made detailed notes in my reflections. She considered video lectures outside class a waste of time and that students can and should get the basic information that was being conveyed in the lectures by reading the text. Without participating in the flipped classroom, this student was adamant that she would not like the method. I noted that she was one of the highest achieving students on campus which in turn caused me to reflect on other interactions with high achieving students in both sections. It occurred to me that most of the positive feedback I received regarding the flipped classroom model was from lower performing students.

4.5.1.4 Theme D: Positive experience with flipped classroom

While I notated frustration with certain students such as those discussed above, there were also students who verbalized how much they enjoyed the process and the video lectures. These seemed to me to be the vast majority. In my reflections, I notated several times how students would come in and tell me how much they enjoyed the lectures. They especially liked how they could pause and rewind if they seemed to miss something or simply not understand it. One student, who was a commuter student living at home, mentioned how she was enjoying the video lectures and that her mother noticed her watching them and started to watch with her. Her mother later visited the campus and stopped by the class to meet me to let me know how much she enjoyed watching the lectures with her daughter.

4.5.2 Semi-structured Focus Groups

The two semi-structured focus groups contained a mix of students from both sections (traditional and flipped). Insights from the focus groups reinforced what I had found in my personal reflections, specifically Themes B, C, and D above. Furthermore, new themes emerged which built upon these previous themes.

4.5.2.1 Theme E: Negative student engagement and student perceptions

Of the nine students who were in the flipped classroom (Section 3) and participated in the focus groups, three reported to have a negative experience and six reported a positive experience. The three students reporting a negative experience were the only students who acknowledged they did not follow the format and often did not watch the video lectures prior to class as assigned. As described in Section 3.5.1, EdPuzzle™ video lectures were the main source of the outside class/behaviorist component of the flipped classroom.

The following are some of the quotes from those who did not like the flipped classroom format from the focus groups.

> **Student 7-Section 3**: It was kind of redundant, in the sense like, I think a lot of us like put off the actual videos we had to watch until a later days, so by the time he would go over the information as if we already watch the videos and then we watch the videos, but we watch them until later, so kind of like I'm already watching them already knowing the information almost, and then probably not watching them as full minute video, and probably sometimes I would skim through them.

> **Student 10**-**Section 3**: I agree with (Student 7-Section 3). I felt like a lot of times in class it was kind of like everyone was just sitting there looking at each other because we hadn't watched the videos, and we hadn't have any questions to ask and the question that we would add, would it be the same questions as we would have had in regular lecture.

> **Student 1-Section 3** (after acknowledging not watching the EdPuzzles™): I did not like it…. Normally in all the classes, or are normally how the structure is, that you went through what the reading was. It (the flipped classroom) was a bit confusing, instead of going over what we have learned we started applying it, so sometimes if you have questions, I'd be like, oh. I'm lost; at least personally.

Upon reflection, it seemed that there were two possible reasons for their dislike of the flipped course format and non-participation. First, there is the possibility that students had negative preconceptions of the flipped format from the outset. Therefore, they did not participate properly which further led them to be confused (as noted by the quote by Student 1-Section 3 above) and/or frustrated. Another possibility is that because they did not properly engage right away in the format, this led to the confusion and/or frustration and then creating a negative perception of the flipped format.

There were two other students who also verbalized they did not like the idea of the flipped classroom, however like the student that I notated in my reflections, these two students were in the traditional classroom and were adamant with their opinions. For example:

> **Student 1-Section 1** (referring to assigned video lectures): I felt like they add an additional stress to my life outside of class. I don't take that long to study, but when you have the video playing for an hour or something there's two videos. I have to do that, for fear of my grade.

While the majority had positive feedback regarding the flipped classroom, the few negative voices each admitted not properly participating in the process. This may have been due to preconceptions regarding the process as seen by those comments from the traditional students.

4.5.2.2 Theme F: Positive student responses aligned to CUCEI scales of innovation and individualization.

The remaining students either verbalized directly that they enjoyed the flipped classroom format, or it was notated in the transcript that they nodded their head in agreement when directly asked. For example, in response to the first question which was aligned with the CUCEI scale of innovation:

> **Student 13-Section 3**: My grades also improved, and I think it's because in class, sometimes it's like, even though I talked a lot in that class sometimes it'd be like scary to repeat like ask a question, but since things were online, I would like to repeat it with no issue. So, if I needed to hear something again, that I could hear it again if I need to take it at my own pace without worrying about the changing slides of just writing down instead of like absorbing the material. It was better than like she said then I can take what I learned there and then ask questions in class.

> **Student 12-Section 3**: I really, really enjoyed the experience because it was so different and I was able to learn a lot more in that way, because it actually forced us to do it on our own, to form our own metacognition. So, how we represented the format was that we were taking it upon ourselves to learn the material and then coming to class, with the questions we didn't understand. So the whole lecture was based on what we didn't understand, and I think a better explanation for it which I really enjoyed, rather than listening to a Professor for the whole 50 minutes and then transcribing everything down. So that's why I enjoyed the approach and I feel like my grades even improved from the first semester to the second semester with that method as well.

> **Student 8-Section -3**: I feel like in classroom we were able to like ask more questions and have more time to go over stuff, because we had already have watched the lecture online, so I understand more.

> **Student 4-Section -3 (responding to the quote above)**: I also felt that, like it was very helpful, like to have like absorb information beforehand seeing him because, like if there was any confusing information you can ask him or email him and class would be more efficiently.

In response to the second focus group question, aligned with the CUCEI scale of Individualization:

> **Student 13-Section 3 (responding that the flipped was preferred)**: …. the only thing was in the beginning I may have struggled but that was on me for maybe not applying myself because you did have to apply yourself to do the work. Cause things are made at a specific time. So, when I actually applied myself and got better at not procrastinating everything then my grade went up, like substantially. Like a whole letter grade.

> **Student 12-Section 3**: I feel like it wouldn't be right to say that just because of the grade. I feel like a lot of work like she said had to do

> with applying ourselves. That's what it was I saw the most beneficial. Yes, was a plus but I think it was very maturing to, to learn that way.

Qualitative data from the semi-structured focus groups seemed to correlate to that of the quantitative data from the CUCEI in relation to two scales, innovation and individualization. The third scale that showed significant difference, involvement, was not expanded upon during the focus groups.

4.5.2.3 Theme G: The Importance of the in-class/constructivist curriculum

As explained in Section 2.4.3.2, the curriculum for in-class activities is intended to be based on a constructivist foundation, one that stresses both active learning and social interaction, ideally emphasizing higher-order concepts. This was done primarily using POGILs, case studies, concept mapping, drawing of anatomical structures and virtual labs. It should be noted that the traditional classroom also participated in many of these active learning activities, however they did them as homework. Therefore, they received the potential benefits of the active learning assignment, but not the social benefits.

Through the focus groups it became clear that some elements of the constructivist curriculum were preferred over others. The questions aligned with the CUCEI scales of innovation and student cohesiveness resulted in very mixed responses, depending on the specific group activities that were discussed. The feedback on case studies was, for the most part, very positive from both sections. For example, from the flipped classroom:

> **Student 12-Section 3**: I think if you can apply it with the outside world and that's when it actually starts to make sense, especially if you want to go into the profession of medicine… Like I think for like the case studies I did find beneficial because you sort of get it forced you to think about what you learned to think about what you already know.

Likewise, from the traditional classroom, Student 9-Section 1. "I liked the case studies. Just because it brought it back to a bigger level more like applicable to health care and I thought it was interesting."

The drawing of anatomical structures and concept maps were also popular with those in the focus groups.

> **Student 13-Section 3**: There was one that I still have, um, that there was this time when we draw a heart and we went step by step and saw different parts of the heart, and that like got it down like that class today did not have any issues with the hard way to ever, and I still haven't memorized. So going through it step by step. Instead of just labeled it, that helped.

However, the views on POGILs were quite mixed. While some liked the POGILs and their connection to applying the information clinically, the majority thought they were too easy and didn’t help expand their understanding. Or, as summarized by Student 11-Section 1, “For me the POGILs were just a bit of tedious.”

What was very clear is that the students did not care for the virtual labs. None of the subjects in the focus groups responded positively to the subject of virtual labs.

> **Student 12-Section 3**: …there was a specific activity that we did online as well. It was like a virtual laboratory simulation. I personally don't like those. I feel like it doesn't help at all, just clicking and moving around certain things to try to explain the point.

> **Student 9-Section 1**: I didn't like it at all, because it would just turn into one person in the group doing the clicking and then you just don't pay attention to what the questions are, or I just didn't get anything out of it, because I didn't study or learn... Yeah, the virtual labs were awful.

In hindsight, after reviewing the qualitative data, I feel I spent too much time on the development of the video lessons and not enough of my energies were put to the in-class/constructivist activities.

4.6 Initial Triangulation of Results

As this is an explanatory sequential mixed-methods study, the qualitiative research is intended to expand and explain the quantitative data. The learning environment data as assessed quantitatively by the CUCEI showed that the flipped classroom reported significantly higher scores with moderate to large effect sizes on three scales. First, the scale of involvement (p = 0.01; r = 0.52), which is associated with the relationship dimension of Moos's scheme, is described as the "extent to which students participate actively and attentively in class discussions and activities" (Fraser et al., 1986). This makes sense as the in-class activities are intended to be group-based and student-centric interactive learning activities (Bishop & Verleger, 2013). The majority reported preferring a flipped classroom format and the in-class activities (Themes D and E). Those who reported preferring the traditional classroom was primarily due to the requirement of watching video lectures outside class and they did not report disliking the in-class activities themselves (except for sometimes being confused because they were not prepared) (Themes B and C).

The next CUCEI scale that was signficantly higher in the flipped classroom was innovation (p = 0.04; r = 0.43). Innovation is associated with Moos's dimension of system maintenance and change. It is described as the "extent to which the instructor plans new, unusual class activities, teaching techniques and assignments" (Fraser et al., 1986). Again, this logically follows in that the flipped classroom itself is a relatively new teaching technique, integrating several other techniques into both the inside class and outside class activities (Arner, 2020). This was acknowledged by the students in the focus group, however as noted above, not all new teaching techniques were seen as a positive (Themes D, E and F).

The last CUCEI scale that indicated signficantly better performance in the flipped classroom was that of individualization (p = 0.03; r = 0.46). Like innovation, individualization is also Moos's dimension of system maintenance and change. It is described as the "extent to which students are allowed to make decisions and are treated differentially according to ability, interest or rate of working" (Fraser et al., 1986). Flipped classrooms emphasize active learning which requires student take ownership of the learning process and to assess their own personal degree of understanding, with activities that involve information gathering, thinking and problem solving (Michael, 2006). Or, as noted previously by Student 12-Section 3, "because it actually forced us to do it on our own, to form our own metacognition" (Themes D and F).

While there was also no significant difference of attitudes toward science between the flipped and traditional classrooms as assessed by the TOSRA, the semi-structured focus groups seemed to point to very different attitudes within each class (Themes E and F). Using the definition of attitudes previously mentioned in Section 2.5.5, "Attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour" (Eagly & Chaiken, 1993, pp. 1–2), some students in the flipped classroom saw the experience favorably, while others reported seeing it unfavorably. This was specfcially evident on how particular students perceived the required lecture videos (Theme C) and certain in-class activities (Theme G). What was interesting to me was that some of the students in the traditional classroom also had these negative views toward required lecture videos even though they were not in the flipped classroom.

There were also no significant differences in the learning outcomes data. However, several students within the flipped classroom who reported they liked and participated in the format (coded as "Engaged(+)") also stated the format increased their confidence ("Confidence(+)") and their grades improved ("Grades (+)") (Theme D). However, those flipped students who admitted to having a negative perception of the flipped classroom format ("Engaged(-)") also reported frustration with the material (Themes C and E). This correlation potentially speaks to how the student's perception of the classroom as well as their attitudes may affect academic performance (see Sections 2.5 and 2.5.5).

4.7 Secondary Quantitative Data Analysis and Results

While the research questions are based on the differences between the traditional (the control) and flipped (the treatment) formats, initial triangulation of quantitative and qualitative data brought to light differences in students' attitudes and perception *within* each classroom as it relates the flipped classroom model and their performance. From my discussion with the high achieving student noted in Section 4.5.1.3 above, who was vehemently against the flipped classroom idea (even though she was not in the flipped classroom), I began to wonder about the potential of certain students to create a bias regarding new instructional modalities, especially if they have been successful in a traditional format. In order to determine if there were any trends of students who reported a positive versus a negative attitude towards the flipped classroom approach and previous academic success, the cumulative grade point average (GPA) of each student in the focus groups was collected by research assistants and aligned with responses in the focus groups.

4.7.1 GPA Data

Students in the semi-structured focus groups were previously coded based on their responses as flipped(+) (liked flip), flipped(-) (did not like flip), or flipped (_) (either neutral or did not express opinion on flip). Those that expressed positive or negative responses to the flipped classroom format were associated with their cumulative grade point average (Cum GPA) prior to the beginning of the semester. This was to see if there was a trend for successful students not to like the flip format (and potentially change in general). All of the flipped classroom students responded with a positive or negative opinion and three traditional students expressed a negative opinion (see Table 4-5). The average GPA of those with a negative opinion was 3.65 while the average GPA for those with a positive opinion was 2.91. While this seemed large, a Mann Whitney U test was run in R and no significant difference was found ($p = 0.05382$). However, it should be noted that of those who responded positively or negatively to the flipped classroom format, those with five of the six highest GPAs responded negatively, including two 4.0 GPA students who were assigned to the traditional classroom.

Table 4-5

Data of positive versus negative feedback on flipped classroom by cumulative GPA prior to course and section enrolled (flipped or traditional)

Positive/Negative Feedback	GPA	Flipped or Traditional
Negative	2.63	Flipped
Negative	3.46	Flipped
Negative	3.85	Flipped
Negative	3.94	Traditional
Negative	4	Traditional
Negative	4	Traditional
Positive	2.71	Flipped
Positive	3.08	Flipped
Positive	3.53	Flipped
Positive	3.09	Flipped
Positive	2.39	Flipped
Positive	2.63	Flipped

There was no significant difference between the cumulative GPA data of those who had positive feedback to the flipped classroom format and those who had negative feedback.

4.8 Addressing the Research Questions

This section will review each of the research questions and findings after the triangulation of quantitative and qualitative data in the previous section.

4.8.1 Research Question 1

How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' perceptions of their learning environment?

Research question 1 is related to the affective domain of learning. From the administration of the CUCEI, it was found that in three of the scales, involvement (associated with Moos' Relationship Dimension), innovation, and individualization (both associated with Moos' System Maintenance and System Change Dimension) students in the flipped classroom scored significantly higher than those in the traditional. Conceptually, this makes sense as the flipped classroom model is a relatively new instructional model that is intended to emphasize active, group-based learning as well as open discussion (Bishop & Verleger, 2013). The quantitative data was supported by the majority of student positive responses gathered qualitatively via my instructor reflections (Theme D) and in the semi-structured focus groups (Themes F and G). What was somewhat surprising to me was not the differences in perceptions between the traditional and flipped, but differences within the flipped classroom itself (Themes E, F and G). These differences in perceptions of the flipped classroom learning environment may be due to students' different preconceptions of what the class should be like. This can be seen from the negative student responses toward the flipped classroom format from students assigned to the traditional class.

4.8.2 Research Question 2

How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' attitudes toward science?

As with research question 1, this research question is also related to the affective domain of learning. Utilizing the TOSRA, the quantitative data regarding student attitudes toward science showed no significant difference between the flipped classroom and that of the treatment group. However, while attitudes regarding science were no different, qualitative data showed that attitudes toward the flipped classroom format varied greatly within the flipped classroom group itself. As noted in Section 2.5.5.2, attitudes are thought to be stored in long-term memory and have three components, a cognitive component, an affective component, and a behavior component (Reid, 2015). As attitudes are in the brain, they must be measured indirectly by assessing behavior. While there was no significant difference found in student attitudes toward science via the TOSRA, the qualitative interviews showed that some students in the traditional classroom had a negative attitude toward the idea of a flipped classroom even though they were not assigned to it. In other words, without experiencing the flipped classroom, they had a negative preconception that they would not like it. As notated in the section above, that may also provide an explanation of why some students in the flipped classroom also had negative perceptions of the format. It was also interesting to note, that of the students who did respond either negatively or positively to the flipped format, the four students with the highest GPAs all had a negative response, regardless of what section they were in.

4.8.3 Research Questions 3 and 4

How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' examination achievement and practical laboratory experience?

Research questions 3 and 4 both refer to student learning outcomes in human anatomy and physiology. Examination scores tended to emphasize the cognitive domain of learning while laboratory practical examinations had both a cognitive and psychomotor component. From the quantitative data, there was no significant difference in learning outcomes between the flipped and traditional classroom groups. This is not surprising as several reviews and meta-analyses have pointed out mixed results with the implementation of the flipped classroom format (Chen et al., 2017; Gillette et al., 2018). However, several students who were coded as having a positive response to the flipped classroom did report their grades improved over the previous semester of anatomy and physiology, and they had increased confidence with the material.

4.9 Summary

This chapter reported the findings of a mixed methods study of a flipped classroom model in an undergraduate human anatomy and physiology class examining students' perceptions of their learning environment, attitude toward science and learning outcomes. To compare the flipped classroom model with a traditional lecture model, quantitative data was initially obtained three ways. First, to examine student perceptions of their environment the CUCEI was utilized (Section 4.4.1). Second, student attitude data was obtained using the TOSRA (Section 4.4.2) and lastly, student outcomes data was gathered using examination and laboratory practical scores (Section 4.4.3). Initial analysis of quantitative data via Mann-Whitney U test and non-parametric test for effect sizes showed significant differences and large effect sizes in three scales of the CUCEI, with the flipped classroom scoring innovation (p = 0.04; r = 0.43), involvement (p = 0.01; r = 0.52), and individualization (p = 0.03; r = 0.46) higher than the traditional class. There were no significant differences found in student attitudes toward science or in learning outcomes.

Qualitative data was gathered through instructor reflections and semi-structured focus groups. The data was entered into Quirkos™ for analysis. Several themes developed from the qualitative data which expanded upon the quantitative results. Specifically, while there were no significant differences between the classes in student attitudes or learning outcomes, within the classrooms, attitudes toward the flipped classroom varied greatly. Specifically, higher performing students seemed to dislike the idea of the flipped classroom format, including those within the traditional class who did not experience it. This was confirmed by collecting additional quantitative data in the form of cumulative GPAs prior the term.

The next chapter will discuss the findings, the limitations of the study and how these findings compare to that of the literature. In addition, it will review the implications of the study for anatomy and physiology instructors as well as future researchers.

Chapter 5. Conclusion

5.1 Introduction

Undergraduate human anatomy and physiology courses are often seen by students as one of the most challenging pre-clinical curricula in both volume and complexity. It is not uncommon for instructors to struggle as well, as they attempt to get across the key concepts to students wishing to enter the medical sciences. Furthermore, as the biomedical sciences continue to expand, there is less allotted time to study anatomy and physiology.

A relatively new instructional model, the flipped classroom, may provide a solution. The flipped classroom model combines two seemingly opposing learning theories: behaviorism, and constructivism, which allows the student to engage with the material in a more applied manner at a deeper level. In many ways, the flipped classroom seems ideal to help anatomy and physiology instructors and students.

However, while research on the flipped classroom has grown, it is still relatively limited, and the results have been mixed. To date, almost 50,000 studies have been located on Google Scholar. Some studies have shown improvements in learning and others show no difference. Furthermore, most of the studies focus solely on learning outcomes, which paints a partial picture by analyzing only the cognitive (and sometimes psychomotor) domain of learning. As the flipped classroom model completely changes how the students interact with the material, the instructor, and each other, a holistic approach must be taken to study the flipped classroom. The approach should be one which takes into account the affective domain of learning. This study researched the flipped classroom approach by examining learning outcomes as well as student perceptions of their classroom environment and their attitudes toward science.

This chapter will summarize the research conducted in this study on the flipped anatomy and physiology classroom. It begins will an overview of the previous chapters, including reviewing the study's objectives, stated research questions, background information on the main topics, methodology, and data collection (Section 5.2). The next section (5.3) will summarize the key findings while Section 5.4 will discuss the limitations of the study. Section 5.5 will examine the study's implications, including recommendations for anatomy and physiology instructors (Section 5.5.1) and future researchers in flipped classroom instruction (Section 5.5.2). Section 5.6 will provide closing remarks summarizing this thesis.

5.2 Overview of the Study

Chapter 1 presented the background of the study. It began by introducing the common challenges facing human anatomy and physiology instructors and students. The chapter then considered the flipped classroom model as a potential solution. This was followed by a discussion of how the study of learning environments could be used in conjunction with traditional learning assessments to create a more holistic picture of the effect of the flipped model on anatomy and physiology classrooms. The conceptual framework of how the flipped classroom model may influence anatomy and physiology student outcomes in all three learning domains (cognitive, psychomotor and affective) was also introduced.

Section 1.4 presented the following purpose statement for the study:

> The purpose of this explanatory sequential mixed methods study is to explore how the implementation of a flipped classroom format impacts the attitudes, perceptions and learning outcomes of traditional undergraduate anatomy and physiology students.

Section 1.5 presented the primary and secondary research questions:

Primary research question: *What is the effect of a flipped classroom environment on undergraduate students in a human anatomy and physiology course?*

Secondary research questions: *How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students'*:

1. *perceptions of their learning environment?*
2. *attitudes toward science?*
3. *examination achievement?*
4. *practical laboratory experience?*

Chapter 2 reviewed the relevant literature related to the study. In Section 2.3 I detailed how researchers and educators consider anatomy and physiology as a cornerstone of the biomedical sciences. I also reviewed the commonly cited traits found within undergraduate anatomy and physiology classrooms and the difficulties which often arise for students and teachers alike. These difficulties are frequently cited as due to the nature of the content (Section 2.3.2.1), improper learning strategies utilized by students (Section 2.3.2.2), and instructors not utilizing appropriate instructional design principles (Section 2.3.2.3).

Next, in Section 2.4, the literature on the flipped classroom model was reviewed. This included a brief history of the method (Section 2.4.1), as well as the common methods for structuring a flipped classroom (Section 2.4.2). Several of the common learning theories cited in the literature to explain the flipped model's potential utility were discussed in Section 2.4.3, and the most recent research on the effectiveness was reviewed in Section 2.4.4.

Section 2.5 presented the literature on classroom learning environment research and student attitudes research. It began with a detailed history of learning environment research (Section 2.5.1). It then detailed the common modalities for assessing classroom learning environments (Section 2.5.2) and reviewed some historically significant and commonly used classroom learning environment instruments (Section 2.5.3). Special emphasis was given to the learning environment instrument used in the study, the CUCEI (Section 2.5.4). The related field of studying student attitudes was discussed in Section 2.5.5, including a discussion on the background, structure, and validation of the instrument used in the study, the TOSRA (Section 2.5.5.2.1).

The methodology utilized in this study is reported in Chapter 3. The philosophical rationale for utilizing a mixed methods approach was discussed (Section 3.3) and a detailed review of all the various components of research design and analysis (Sections 3.4-3.8).

Chapter 4 reported the data analysis and the results of the study. The results from the initial quantitative data collection, including CUCEI, TOSRA and learning outcomes data, were reviewed in Section 4.4. This was then followed by a discussion of the qualitative data that was collected to expand upon the previously collected quantitative data (Section 4.5). Section 4.7 explained how, after both sources of data were integrated through triangulation (Section 4.6), and more quantitative data in the form of student GPA data was found to be needed. The addition of this data was again integrated via triangulation, and the individual research questions were addressed (Section 4.8).

As described above (Section 5.1), this final chapter will summarize the study's findings, the study's limitations, and the implications of the study results. This includes recommendations for practice and future research, and my final conclusions (Sections 5.3-5.5).

5.3 Summary of Key Findings

The findings of Chapter 4 are examined in the following three subsections dedicated to the research questions and discusses the results in relation to previous research.

5.3.1 Research Question 1: How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' perceptions of their learning environment?

The quantitative data show that there are differences in the students' perceptions of their flipped learning environment compared to those in the traditional class. The CUCEI was administered to assess anatomy and physiology students' perceptions of their learning environment. The mean scale scores of the flipped classroom showed significantly higher responses with moderate-to-large effect sizes than that of the traditional classroom on three scales of the CUCEI: innovation ($p = 0.04$; $r = 0.43$), involvement ($p = 0.01$; $r = 0.52$), and individualization ($p = 0.03$; $r = 0.46$). These results seem logical due to the nature of the flipped classroom. Specifically, the scale of innovation makes sense as the flipped classroom is a relatively new instructional model, so it is highly likely that most students would not have been introduced to it previously and, hence, found it innovative.

Furthermore, the flipped classroom format allows for more active learning modalities within the classroom, which also may be seen as innovative. The active learning modalities may also explain why involvement and individualization were higher as well. Active learning is intended to be student-centric and is often group-based (Michael, 2006). As active learning forces students to reflect upon their ideas and create an environment engaging them in information gathering and problem-solving, it is no surprise that students in the flipped classroom score these three scales higher.

While there have not been previous studies analyzing the learning environment in a flipped anatomy and physiology classroom, there have been a few studies that examined the learning environment in undergraduate STEM courses. Several of results from these studies were consistent with my results. For example, Clark et al. (2014) utilized the CUCEI on an industrial engineering course on facility layout and material handling. In administering the CUCEI before and after implementing the flipped model, also found a significant difference in the same three scales of individualization, innovation, and involvement. In addition, they also found students ranked personalization higher after the flip (R. M. Clark et al., 2014). Additionally, McNally et al. (2017) sampled undergraduate and postgraduate health science students who underwent the flipped model. In order to assess the students' perceptions of the flipped classroom learning environment, a modified version of the CUCEI was used to compare those who endorsed the flip model to those students who resisted it. When asked about their ideal classroom, the flipped endorsers rated the scales of involvement, innovation, and cooperation significantly higher than the flipped resisters (note: individualization was not part of the revised version of the CUCEI used in the study) (McNally et al., 2017). From the above studies and the current research, it seems that students who find the components of individualization, innovation, and involvement important are likely to have a positive response to the flipped classroom model.

The qualitative data gathered from instructor reflections and semi-structured focus groups reinforced the data gathered from the CUCEI. Except for a vocal minority (which will be discussed in the next section), the majority of feedback noted, throughout the term, in my instructor reflections journal was positive, with students liking the in-class activities as well as the outside class lectures. This is similar to what both McNally et al. (2017) and He, Holton, Farkas, & Warschauer (2016) found, as certain students resisted the idea of the flipped classroom while others embraced it. The students who were favorable to the flipped model seemed to enjoy that classes were more interactive, and several told me how much they liked being able to stop and rewind/re-watch the lectures, a benefit also noted by Clark et al.

(2014). Even students in the semi-structured focus groups who reported not liking the flipped format said the reason for their animosity to the format was due to having to watch videos outside of class. This is a common complaint, also cited by He et al. (2016) and Şengel (2016). Most students actually admitted liking the active learning components in class.

5.3.2 Research Question 2: How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' attitudes toward science?

The initial quantitative data collected from the TOSRA to assess students' attitudes toward science showed no significant difference between the flipped and traditional classrooms. This was reinforced in the qualitative data, as none of the student feedback appeared to show any differences between the two classrooms judged by Klopfer's classification scheme (as explained in Section 2.5.5.3).

However, while attitudes toward science did not differ between classes, attitudes toward the flipped classroom model varied greatly within each class. As mentioned in the above section, several students from both classes seem to have negative attitudes toward the flipped classroom model. While the negative responses for those assigned to the flipped class may be their genuine responses to experiencing the model, some from the traditional class were shown to have a negative preconception of what the flipped class would be like. This is often found in the literature, as cited above with McNally et al. (2017) labeling those who embraced the flipped classroom "flipped endorsers" and those who did not endorse the pre-learning aspect "flipped resisters". When attempting to differentiate those who tended to resist versus endorse, I examined GPA data to see if high-performing students differed from lower-performing students. While not statistically significant ($p = 0.053$), the fact that all four of the highest GPA students responded negatively does show a potential trend that should be examined with a larger sample size. Another potential variable to be examined in the future that may contribute to differences in openness to the flipped classroom would be student major. Anderton, Shan Chiu, et al. (2016) found that biological science students were least open to progressive teaching modalities in anatomy and physiology, while health and sport-related majors were more open. So, while the data showed no difference in student attitudes toward science, qualitative data did provide evidence that students have very different attitudes toward the flipped method, whether assigned to the class or not.

5.3.3 Research Questions 3 and 4: How do flipped classrooms in anatomy and physiology differ from traditional lecture formats in terms of students' examination achievement, and practical laboratory experience?

In examining the quantitative data comparing examination and practical laboratory performance between the flipped and traditional classrooms, no significant difference was found. This is not surprising as reviews of the flipped classroom model point out that while many studies report improvements in student learning, several others show no significant difference (Akçayır & Akçayır, 2018; Chen et al., 2017; O'Flaherty & Phillips, 2015). A potential reason that there was no significant difference in learning outcomes is that part of the flipped classroom approach's suggested benefit is active learning modalities used in the classroom. In this study, the traditional students were given the same active learning assignments (e.g., POGILs, case studies, virtual labs) as homework. Therefore, while the traditional class missed the social component, they still had assignments that were intended to increase critical thinking. DeLozier & Rhodes (2017) as well as J. L. Jensen, Kummer, & Godoy (2015), argue that it is the value of these activities that is important, whether the setting is traditional/lecture-based or flipped. So, the key for any success in the flipped classroom may not be flipping the classroom itself, but rather the active learning modalities assigned.

Several students in the focus groups reported that their grades went up and had increased confidence with the material due to the flipped classroom's implementation. When cross-referenced with cumulative GPA and how they were coded in relation to liking/disliking the flip, these students all were lower-performing students who reported liking the flipped format. These are similar findings to those found in Day's (2018) study of a flipped gross anatomy class. In that study, improvements in knowledge acquisition and transfer happened to a greater degree in lower-performing students than higher-performing students (Day, 2018). Therefore, student success in the flipped classroom model may be determined in part by their openness to modalities, which may be related to their previous academic achievement.

5.4 Limitations to the Study

This quasi-experimental, explanatory mixed methods design had a small convenience sample (i.e., the subjects were the researcher's students), which makes it difficult to detect significant differences and to generalize the results. Each section was unusually small as most sections of the class I taught in the past consisted of twice as many students. I considered attempting to collect more data by replicating the study the following year when the course was next offered, but unfortunately due to the COVID-19 pandemic this was not an option.

This increases the risk of bias (in both the researcher's interpretations and the subjects' responses) and limits generalization (L. Cohen et al., 2018). To decrease subject bias and the perception of any teacher coercion, it was made clear that participation in the study had no impact on the students' grades and that their responses would be kept confidential. Random numbers were assigned, and other faculty members administered the surveys and semi-structured focus groups. Transcription of focus groups was created by research assistants, who also helped align some quantitative data elements (e.g., cumulative GPA) with qualitative responses. However, it is possible that students' may have still feared that I may somehow determine their specific responses, which may have kept them from being completely candid and truthful on the surveys or in the focus groups.

A potential threat to internal validity is that of subject mortality (i.e., dropout) (Fraenkel, Wallen, & Hyun, 2012). The semi-structured focus groups were conducted the following fall semester, after a summer break. This was intended to allow time for students to reflect on their experiences. Unfortunately, due to changes at the university, a program (athletic training) was dropped, and several students who were subjects in the study transferred to other universities and were unwilling to participate in focus groups (either in-person or virtually). This decreased the sample size.

The small sample size of students at one university is the greatest threat to external validity/generalization. However, my goal here was to report what happened in my specific classroom when exposing one class to a particular form of the flipped classroom method. There are many variations to the flipped classroom in design, development, and implementation, so a more extensive study involving many teachers from many schools may not give the same specific results because each teacher would implement the flipped format in their own way. In addition, many studies use single-group study designs (Bishop & Verleger, 2013). If outcomes are assessed, they are often compared to previous classes or coursework with a completely different curriculum. With this small sample, utilizing a traditional class for a control, I was able to achieve very specific results. Both classes received the same content, so the primary difference was the order/time of content delivery. This allowed the flipped classroom to get behaviorist/lecture-based material outside class and constructivist/active learning material inside class with the traditional classroom being vice versa. So other than a temporal change, the primary difference was the lack of constructivist interactions between students and between student and teacher.

As a result of the small sample size, I could not do a factor analysis on the CUCEI and TOSRA data. However, these two instruments have been well validated in past studies. I was still able to show statistically significant differences between the flipped and traditional classrooms on three scales of the CUCEI. In addition, the qualitative data added context to the quantitative data, allowing students' voices to be heard. Together, they show distinct differences in how students view the flipped classroom and how these differences seem to be in place regardless of participation in the flipped format. So even with the small sample size, the specific results can still provide insights for future research as well as the implementation of the flipped classroom model in an anatomy and physiology classroom.

5.5 Implications of this Study

As noted by several authors (Arner, 2020; Betihavas et al., 2016; Burke & Fedorek, 2017) multiple considerations must be taken into account when flipping a classroom, each of which can contribute to the success or failure of the implementation of the flip. In addition to the data described above, it was a review of these previous findings that led me to reflect on the development, launch and implementation of this study and how certain changes may produce different outcomes in the future. These insights pertain to both future implementations of the flipped classroom by anatomy and physiology instructors as well as researchers interested in the flipped model. These are reviewed in the following subsections.

5.5.1 Implications for Anatomy and Physiology Instructors Interested in Flipping their Class

This section will review some of my reflections pertaining to the data collected and research pertaining to flipped classrooms.

5.5.1.1 Design and Development

When deciding on whether to flip a class, instructors should have clear outcome goals in mind. As an anatomy and physiology instructor, my goal was to help students who are often challenged by the volume and complexity of information in a typical A&P classroom prepare for clinical coursework in the future. As my data demonstrated, as well as that of previous research, flipped classrooms can positively affect the affective domain creating changes to the learning environment (R. M. Clark et al., 2014; McNally et al., 2017). When examining potential changes in the cognitive domain in the flipped classroom by studying learning outcomes, the research is inconsistent (Betihavas et al., 2016; Bishop & Verleger, 2013), tending to be slightly better in the flipped class or, as in the case of this study, neutral. The difference between flipped classrooms with improvement in student outcomes is thought to be the proper development of the curriculum. Therefore, if an anatomy physiology instructor wishes to flip their class with traditional academic goals in mind, careful consideration must be taken in the class's design and development.

While there is no set structure for the flipped classroom model, the basic structure of outside class content and higher-order thinking inside class should be designed so there is thoughtful alignment between the two (Arner, 2020; O'Flaherty & Phillips, 2015; O'Shea, 2020a). This is an area where I could have done better, as to ensure the behaviorist curriculum as identical between classes, I focused perhaps too much on the outside class content. Ideally, it is inside the class where active learning takes place in a constructivist manner. While I chose assignments and activities that I thought would align well with the lectures, the students' feedback was quite mixed.

In addition to anatomy and physiology instructors needing to carefully contemplate how the course is structured, they must also consider the time and resources required to create a successful class. In my instructor reflections journal (detailed in Section 4.5.1), I recount the difficulties I had creating the flipped classroom. What was especially time-consuming was the creation of the video lessons. This is not uncommon, as O'Flaherty and Phillips (2015) noted in their scoping review, the "lead in time for faculty was intense" (p. 89). Some advocate using pre-made video content rather than their own, with Arner (2020) suggesting that the outside class content's origin is less important than the quality of design and instruction. To this I agree, but when specifically asked in the focus groups if they preferred having videos from the instructor or others on YouTube, it was unanimous that they preferred instructor-made videos (although several admitted watching other videos to complement the content). As mentioned above, I created the video lessons in such a way to make the lecture-based/behaviorist content as similar as possible. This allowed me to control for the lectures, however, Walker et al. (2020) suggest that this is not ideal. Rather they endorsed that, "video lectures should be designed more as a guide for reading and increasing curiosity rather than as explanations of everything stated in the literature" (Walker et al., 2020, p. 8).

5.5.1.2 Launch

One of the most startling realizations that occurred to me when analyzing the data was the stark differences in student attitudes toward the flipped method and their respective perception of the classroom learning environment. While some students reported enjoying the format, the level of animosity of a minority of others toward the flipped format was relatively high. This seems common in the literature, as often students are not prepared for the transition into the flipped classroom and, as a result, do not properly watch the lectures at home (Burke & Fedorek, 2017). Several authors discuss the importance of student "buy-in" (Burke & Fedorek, 2017; DeRuisseau, 2016; Gilboy et al., 2015; Heyborne & Perrett, 2016). Betihavas et al. (2016) stated in their systematic review of the flipped classroom in nursing education that student engagement in the flipped classroom format was accomplished when instructors properly informed and rationalized the purpose of the flipped classroom model to the students. So, it seems that educating students about the potential benefits of the flipped classroom model at the beginning is essential to the "buy-in" that has eluded past studies.

5.5.1.3 Implementation

The nature of the flipped classroom enables a greater dialogue between instructor and students (McLean et al., 2016). This needs to be maintained by the instructor, and while it may be tempting to fall back into old methods and become the “sage on the stage,” the instructor must be continually aware and remember their changing role. While I would sometimes begin class with a “mini-lecture” to review some key concepts from the previous online lectures, the students who would not participate by watching the online lectures would sometimes coerce me into double-lecturing. Perksy and McLaughlin (2017) warn how this should be avoided. For students who did prepare, class time would become redundant. So it is important to keep class time for constructivist/active learning activities, providing feedback and giving continuous assessment (Zainuddin & Perera, 2019).

5.5.2 Implications for Future Research

This study added to the increasing amount of research on the relatively new instructional format of the flipped classroom conducted specifically within an undergraduate anatomy and physiology classroom. By examining the students’ perceptions of their learning environment, their attitudes in the class, and their performance on assessments, this study was able to look at the flipped classroom in terms of all three domains of learning: cognitive, psychomotor, and affective. The following recommendations for future research are suggested based on the results of this study.

One of the benefits that was unique to this study is that the nature of the flipped classroom was tightly controlled so that students in both the control and treatment groups were exposed to the same curriculum. In this way, the primary differences were temporal (i.e., did they receive the lectures before class or in class) and social (i.e., were they exposed to active learning in the classroom with interactions with peers and the instructor, or did they have to do it as homework). However, if this could be expanded and replicated with a larger sample from multiple classrooms, it would allow for a greater generalization of findings. Further study is recommended where a two-group, pre/post design is utilized and the flipped online curricula emulate that the traditional curricula, but in reverse order and with larger sample sizes.

In addition, it was noted in this study that the highest performing students were all “flip resisters” (to use the term coined by McNally et al. (2017)). A special note was made of those successful students who were not assigned to the flipped classroom yet were contrary to the idea when discussed in the semi-structured focus groups. Previous research also found that students of specific majors resisted the flipped classroom model more than other majors (Anderton, Shan Chiu, et al., 2016). This implies that certain students have certain preconceptions regarding the flipped classroom. Future research can examine the common traits of students who do not like the idea of the flipped classroom, the common objections, and the best ways for instructors to overcome these objections.

This study was also unique because it used well-established instruments to assess the perceived learning environment and student attitudes in addition to assessing learning outcomes. This is unique for several reasons. First, it allowed the study to give a holistic view of a flipped undergraduate anatomy and physiology classroom. It did this by assessing all three learning domains (cognitive, psychomotor, and affective). Second, little research has been done on students' perceived learning environment within the flipped classroom at the tertiary level. Furthermore, most learning environment research has been done at the primary and secondary levels, and there has been a call for more research like this examining classroom environment at the tertiary level (Alansari & Rubie-Davies, 2020). Third, while the assessment of student attitudes is common when studying the flipped classroom, few use well-established surveys, and instead, the researchers often create the instruments. Reid (2015) points out that such student attitudes studies are often methodologically questionable and cannot give the answers we need. Future studies of the flipped classroom should assess student attitudes using well-validated assessments such as the TOSRA and follow the recommendations set by Reid for properly assessing student attitudes (Reid, 2015, p. 37). Future research on flipped classrooms should continue to incorporate learning outcomes, student attitudes, and learning environment to help give a holistic view of what is happening in the classroom.

5.6 Final Summation

Anatomy and physiology instructors are confronted with the task of preparing students with the foundational knowledge they need for future clinical coursework. This is especially challenging due to the volume and complexity of the subject.

Furthermore, applying the concepts in a clinical setting requires higher-order thinking. With decreased time allotted to A&P courses and poor student learning strategies, some see new health professionals as not adequately prepared for medical settings (J. P. Collins, 2009; Singh et al., 2019; Smith et al., 2014).

This thesis examined the flipped classroom instructional model which may solve some of the common problems faced by A&P instructors and students. As the flipped classroom model completely changes the instructor-student dynamic inside and outside the classroom, the study took a holistic view by examining all three domains of learning: cognitive, psychomotor, and affective. This was done by assessing the students' perspectives on their learning environment, attitudes, and learning outcomes.

As a result of this study, I conclude that the implementation of a flipped classroom format does impact the attitudes and perceptions of undergraduate anatomy and physiology students. Specifically, I conclude that the flipped classroom increases students' perception of 1) their active participation within the classroom (based on the CUCEI scale of Involvement), 2) the extent the instructor is innovative in their teaching techniques (based on the CUCEI scale of Innovation), and 3) their ability to take control of their own learning and have it customized to their own ability (based on the CUCEI scale of Individualization). I conclude that students' positive or negative preconceptions of the flipped format influence how they approach the material, which is a significant determinant of the class's success. Lastly, I also conclude that the flipped classroom format shows promise for improving learning outcomes, especially for lower-performing anatomy and physiology students. However, thoughtful design and implementation of the flipped method are essential for success.

Anatomy and physiology instructors who are interested in implementing the method can take lessons from this study. There is no single method for designing a flipped classroom, and as technology changes, the flipped classroom will change as well. That said, whatever tools are used, the curriculum must be well planned, with continuous assessment and feedback (Arner, 2020; Zainuddin & Perera, 2019). Most importantly, student buy-in is critical. While this itself can be challenging, according to Betihavas et al. (2016), student engagement is best achieved by clearly communicating the rationale and purpose of the flipped classroom method. Or, put another way, to get the most out of the flipped classroom design, A&P instructors need to clearly explain to the students the benefits of the flipped format to their knowledge acquisition and future goals.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research and Development*, *34*(1), 1–14. https://doi.org/10.1080/07294360.2014.934336

Afari, E., Aldridge, J. M., & Fraser, B. J. (2012). Effectiveness of using games in tertiary-level mathematics classrooms. *International Journal of Science and Mathematics Education*, *10*(6), 1369–1392. https://doi.org/10.1007/s10763-012-9340-5

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers and Education*. https://doi.org/10.1016/j.compedu.2018.07.021

Alansari, M., & Rubie-Davies, C. (2020, April 1). What about the tertiary climate? Reflecting on five decades of class climate research. *Learning Environments Research*, Vol. 23. https://doi.org/10.1007/s10984-019-09288-9

Aldridge, J. M., Dorman, J., & Fraser, B. J. (2004). Use of Multitrait-Multimethod Modelling to Validate Actual and Preferred Forms of theTechnology-Rich Outcomes-Focused Learning Environment Inventory (Troflei). *Australian Journal of Educational & Developmental Psychology*, *4*. Retrieved from http://www.newcastle.edu.au/journal/ajedp/

Aldridge, J. M., Fraser, B. J., Bell, L., & Dorman, J. (2012). Using a New Learning Environment Questionnaire for Reflection in Teacher Action Research. *Journal of Science Teacher Education*, *23*(3), 259–290. https://doi.org/10.1007/s10972-012-9268-1

Anderton, R. S., Chiu, L., Aulfrey, S., Anderton, R. S., Shan Chiu, L., & Aulfrey, S. (2016). Student Perceptions to Teaching Undergraduate Anatomy in Health Sciences. *International Journal of Higher Education*, *5*(3), 201–216. https://doi.org/10.5430/ijhe.v5n3p201

Anderton, R. S., Shan Chiu, L., Aulfrey, S., Chiu, L., Aulfrey, S., Anderton, R. S., … Aulfrey, S. (2016). Student Perceptions to Teaching Undergraduate Anatomy in Health Sciences. *International Journal of Higher Education*, *5*(3), 201–216. https://doi.org/10.5430/ijhe.v5n3p201

Arner, T. (2020). Instructional Design: Evidence-Based Practices in the Flipped Classroom. In Z. Walker, D. Tan, & N. K. Koh (Eds.), *Flipped Classrooms with Diverse Learners* (pp. 17–34). Springer Singapore.

Aubrey, K., & Riley, A. (2019). *Understanding & Using Educational Theories* (2nd ed.; J. Clark, Ed.). London: SAGE Publications Ltd.

Bergmann, J., & Sams, A. (2012). *Flip YOUR Classroom: Reach Every Student in Every Class Every Day*. International Society for Technology in Education.

Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for 'flipping out': A systematic review of the flipped classroom in nursing education. *Nurse Education Today*, *38*, 15–21. https://doi.org/10.1016/j.nedt.2015.12.010

Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. *ASEE Annual Conference and Exposition, Conference Proceedings*. https://doi.org/10.18260/1-2--22585

Brown, P. J. P. (2010). Process-oriented guided-inquiry learning in an introductory anatomy and physiology course with a diverse student population. *American Journal of Physiology - Advances in Physiology Education*, *34*(3), 150–155. https://doi.org/10.1152/advan.00055.2010

Brown, S. J., White, S., & Power, N. (2017). Introductory anatomy and physiology in an undergraduate nursing curriculum. *Adv Physiol Educ*, *41*, 56–61. https://doi.org/10.1152/advan.00112.2016.-Using

Buckley, S., Coleman, J., Davison, I., Khan, K. S., Zamora, J., Malick, S., … Sayers, J. (2009). The educational effects of portfolios on undergraduate student learning: A Best Evidence Medical Education (BEME) systematic review. BEME Guide No. 11. *Medical Teacher*, Vol. 31, pp. 282–298.

https://doi.org/10.1080/01421590902889897

Burke, A. S., & Fedorek, B. (2017). Does “flipping” promote engagement?: A comparison of a traditional, online, and flipped class. *Active Learning in Higher Education*, *18*(1), 11–24. https://doi.org/10.1177/1469787417693487

Chen, F., Lui, A. M., & Martinelli, S. M. (2017). A systematic review of the effectiveness of flipped classrooms in medical education. *Medical Education*, *51*(6), 585–597. https://doi.org/10.1111/medu.13272

Clark, R. M., Kaw, A., & Besterfield-Sacre, M. (2016). Comparing the effectiveness of blended, semi-flipped, and flipped formats in an engineering numerical methods course. *Advances in Engineering Education*, *5*(3).

Clark, R. M., Norman, B. A., & Besterfield-Sacre, M. (2014). Preliminary experiences with “Flipping” a facility layout/material handling course. In H. Guan, Y., Liao (Ed.), *IIE Annual Conference and Expo 2014* (pp. 1194–1202). Montreal, Canada.

Cliff, W. H., & Wright, A. W. (1996). Directed Case Study Method for Teaching Human Anatomy and Physiology. *Advances in Physiology Education*, *15*(1), 19–28. https://doi.org/10.1017/CBO9781107415324.004

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). https://doi.org/10.4324/9780203771587

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). New York, NY, NY: Routledge.

Collins, J. P. (2009). Are the changes in anatomy teaching compromising patient care? *Clinical Teacher*, *6*(1), 18–21. https://doi.org/10.1111/j.1743-498X.2008.00256.x

Collins, T. J., Given, R. L., Hulsebosch, C. E., & Miller, B. T. (1994). Status of gross anatomy in the U.S. and Canada: Dilemma for the 21st century. *Clinical Anatomy*, *7*(5), 275–296. https://doi.org/10.1002/ca.980070509

Creswell, J. W. (2009). *Reasearch Design: Qualitative, Quantitative and Mixed*

Methods Approches (3ed.). Thousand Oaks, CA, USA: SAGE Publications.

Creswell, J. W. (2014). *A Concise Introduction to Mixed Methods Reasearch*. Thousand Oaks, CA , USA: SAGE Publications Inc.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and Conducting Mixed Methods Research* (3rd ed.). SAGE Publications.

Crotty, M. (1998). *The Foundations of Social Research: Meaning and perspective in the research process*. Thousand Oaks, CA, USA: SAGE Publications.

Daniel, B. K., & Harland, T. (2018). *Higher Education Research Methodology*. New York, NY: Routledge.

Day, L. J. (2018). A gross anatomy flipped classroom effects performance, retention, and higher-level thinking in lower performing students. *Anatomical Sciences Education*, *11*(6), 565–574. https://doi.org/10.1002/ase.1772

DeLozier, S. J., & Rhodes, M. G. (2017, March 1). Flipped Classrooms: a Review of Key Ideas and Recommendations for Practice. *Educational Psychology Review*, Vol. 29, pp. 141–151. https://doi.org/10.1007/s10648-015-9356-9

DeRuisseau, L. R. (2016). The flipped classroom allows for more class time devoted to critical thinking. *Advances in Physiology Education*, *40*(4), 522–528. https://doi.org/10.1152/ADVAN.00033.2016

Dobson, J. L. (2013). Retrieval practice is an efficient method of enhancing the retention of anatomy and physiology information. *American Journal of Physiology - Advances in Physiology Education*, *37*(2), 184–191. https://doi.org/10.1152/advan.00174.2012

Eagly, A. H., & Chaiken, S. (1993). The psychology of attitudes. In *The psychology of attitudes.* Orlando, FL, US: Harcourt Brace Jovanovich College Publishers.

El-Sayed, R. E.-S. H., & El-Sayed, S. E.-H. A. E.-R. (2013). Video-based lectures: An emerging paradigm for teaching human anatomy and physiology to student nurses. *Alexandria Journal of Medicine*, *49*(3), 215–222. https://doi.org/10.1016/j.ajme.2012.11.002

Ellis, H. (2001). Teaching in the dissecting room. *Clinical Anatomy*, *14*(2), 149–151. https://doi.org/10.1002/1098-2353(200103)14:2<149::AID-CA1023>3.0.CO;2-U

Ervin, J. A. (2018). Teaching science to non-science majors: An action research study using a learner-centered, inquiry-based approach. *Journal of Ethnographic & Qualitative Research*, *13*, 79–91.

Eseonu, O., Carachi, R., & Brindley, N. (2013). Case-based anatomy teaching: A viable alternative? *Clinical Teacher*, *10*(4), 236–241. https://doi.org/10.1111/tct.12013

Espinoza, F. (2012). *The Nature of Science*. Lanham, Maryland: Rowmand & Littlefield Publishers Inc.

Evans, L., Bosch, M. L. V., Harrington, S., Schoofs, N., & Coviak, C. (2019). Flipping the Classroom in Health Care Higher Education: A Systematic Review. *Nurse Educator*, *44*(2), 74–78. https://doi.org/10.1097/NNE.0000000000000554

Ferrer-Torregrosa, J., Jiménez-Rodríguez, M. Á., Torralba-Estelles, J., Garzón-Farinós, F., Pérez-Bermejo, M., & Fernández-Ehrling, N. (2016). Distance learning ects and flipped classroom in the anatomy learning: Comparative study of the use of augmented reality, video and notes. *BMC Medical Education*, *16*(1), 1–9. https://doi.org/10.1186/s12909-016-0757-3

Foertsch, J., Moses, G., Strikwerda, J., & Litzkow, M. (2002). Reversing the lecture/homework paradigm using eTEACH® web-based streaming video software. *Journal of Engineering Education*, *91*(3), 267–274. https://doi.org/10.1002/j.2168-9830.2002.tb00703.x

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education (8th ed.).* New York, NY.

Fraser, B. J. (1977). Selection and validation of attitude scales for curriculum evaluation. *Science Education*, *61*(3), 317–329. https://doi.org/10.1002/sce.3730610307

Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education*, *62*(4), 509–515. https://doi.org/10.1002/sce.3730620411

Fraser, B. J. (1998). Classroom Environment Instruments: Development, Validity And Applications. *Learning Environments Research*, *1*(1), 7–34. https://doi.org/10.1023/a:1009932514731

Fraser, B. J. (2012a). Classroom Environment. In *Classroom Environment* (Vol. 234). https://doi.org/10.4324/9780203125885

Fraser, B. J. (2012b). Classroom Learning Environments: Retrospect, Context and Prospect. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 1191–1239). https://doi.org/10.1007/978-1-4020-9041-7

Fraser, B. J. (2014). Classroom Learning Environments Historical and Contemporary Perspectives. In *Handbook of research on science education* (Vol. 2, pp. 104–119). Retrieved from https://espace.curtin.edu.au/handle/20.500.11937/23061

Fraser, B. J. (2019). Milestones in the Evolution of the Learning Environments Field over the Past Three Decades. In *Thirty Years of Learning Environments*. https://doi.org/10.1163/9789004387720

Fraser, B. J., Aldridge, J. M., & Adolphe, F. S. G. (2010). A cross-national study of secondary science classroom environments in Australia and Indonesia. *Research in Science Education*, *40*(4), 551–571. https://doi.org/10.1007/s11165-009-9133-1

Fraser, B. J., Anderson, G. J., & Walberg, H. J. (1982). *Assessment of Learning Environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI). Third Version*. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&NEWS=N&PAGE=fulltext&D=eric2&AN=ED223649

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1995). Evolution and validation of a personal form of an instrument for assessing science laboratory classroom environments. *Journal of Research in Science Teaching*, *32*(4), 399–422.

https://doi.org/10.1002/tea.3660320408

Fraser, B. J., & Lee, S. U. (2015). USE OF TEST OF SCIENCE RELATED ATTITUDES (TOSRA) IN KOREA. In M. S. Khine (Ed.), *Attitude measurements in science education : classic and contemporary approaches* (pp. 293–308). Charlotte, NC: Information Age Publishing.

Fraser, B. J., Treagust, D. F., & Dennis, N. C. (1986). Development of an instrument for assessing classroom psychosocial environment at universities and colleges. *Studies in Higher Education*, *11*(1), 43–54. https://doi.org/10.1080/03075078612331378451

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior*, *47*(1), 109–114. https://doi.org/10.1016/j.jneb.2014.08.008

Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., & Broedel-Zaugg, K. (2018). A meta-analysis of outcomes comparing flipped classroom and lecture. *American Journal of Pharmaceutical Education*, *82*(5), 433–440. https://doi.org/10.5688/ajpe6898

Glass, D. J. (2014). *Experimental Design for Biologists* (2nd ed.). New York, NY: Cold Springs Harbor Laboratory Press.

Griff, E. R. (2016). Changing undergraduate human anatomy and physiology laboratories: Perspectives from a large-enrollment course. *Advances in Physiology Education*, *40*(3), 388–392. https://doi.org/10.1152/advan.00057.2016

Grix, J. (2002). Introducing Students to the Generic Terminology of Social Research. *Politics*, *22*(3), 175–186. https://doi.org/10.1111/1467-9256.00173

Guest, G., MacQueen, K., & Namey, E. (2011). *Applied thematic analysis*. Retrieved from https://books.google.com/books?hl=en&lr=&id=Hr11DwAAQBAJ&oi=fnd&pg=PP1&dq=applied+thematic+analysis+guest&ots=Xi5FvMDKvB&sig=KBLY6WLhH3A2qWX2gpYYBQ_F3Q4

Gultice, A., Witham, A., & Kallmeyer, R. (2015). How We Teach: Classroom And Laboratory Research Projects Are your students ready for anatomy and physiology? Developing tools to identify students at risk for failure. *Adv Physiol Educ*, *39*, 108–115. https://doi.org/10.1152/advan.00112.2014.-High

Guyer, P. (2010). The Cambridge companion to Kant's critique of Pure Reason. In *The Cambridge Companion to Kant's Critique of Pure Reason*. https://doi.org/10.1017/CCOL9780521883863

Hasan, A., & Fraser, B. J. (2015). Effectiveness of teaching strategies for engaging adults who experienced childhood difficulties in learning mathematics. *Learning Environments Research*, *18*(1), 1–13. https://doi.org/10.1007/s10984-013-9154-6

Hase, H. D., & Goldberg, L. R. (1967). Comparative Validity of Different Strategies of Constructing Personality Inventory Scales. *Psychological Bulletin*, *67*(4), 231–248. https://doi.org/10.1037/h0024421

He, W., Holton, A., Farkas, G., & Warschauer, M. (2016). The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions. *Learning and Instruction*, *45*, 61–71. https://doi.org/10.1016/j.learninstruc.2016.07.001

Herreid, C. F., & Schiller, N. A. (2013). Case Studies and the Flipped Classroom. *Science*, *42*(5), 62–66. https://doi.org/10.1126/science.139.3555.660-a

Heyborne, W., & Perrett, J. (2016). To Flip or Not to Flip? Analysis of a Flipped Classroom Pedagogy in a General Biology Course. *Journal of College Science Teaching*, *045*(04), 31. https://doi.org/10.2505/4/jcst16_045_04_31

Higgins-Opitz, S. B., & Tufts, M. (2015). Performance of first-year health sciences students in a large, diverse, multidisciplinary, first-semester, physiology service module. *Advances in Physiology Education*, *38*(2), 161–169. https://doi.org/10.1152/advan.00067.2013

Hoidn, S. (2017). *Student-Centered Learning Environments in Higher Education Classrooms*. Springer Nature.

Jensen, J. L., Holt, E. A., Sowards, J. B., Heath Ogden, T., & West, R. E. (2018). Investigating Strategies for Pre-Class Content Learning in a Flipped Classroom. *Journal of Science Education and Technology*, *27*(6). https://doi.org/10.1007/s10956-018-9740-6

Jensen, J. L., Kummer, T. A., & Godoy, P. D. M. (2015). Improvements from a Flipped Classroom May Simply Be the Fruits of Active Learning. *Life Science Education*, *14*(2013), 1–12. https://doi.org/10.1187/10.1187/cbe.14-08-0129

Jensen, M. (2014). *POGIL Activities for Introductory Anatomy and Physiology Courses*. Lancaster, PA: Wiley.

Johnston, A. N. B., Hamill, J., Barton, M. J., Baldwin, S., Percival, J., Williams-Pritchard, G., … Todorovic, M. (2015). Student learning styles in anatomy and physiology courses: Meeting the needs of nursing students. *Nurse Education in Practice*, *15*(6), 415–420. https://doi.org/10.1016/j.nepr.2015.05.001

Khalil, M. K., & Elkhider, I. A. (2016). Applying learning theories and instructional design models for effective instruction. *Advances in Physiology Education*, *40*(2), 147–156. https://doi.org/10.1152/advan.00138.2015

Khanova, J., Roth, M. T., Rodgers, J. E., & McLaughlin, J. E. (2015). Student experiences across multiple flipped courses in a single curriculum.: Still OneSearch - A.T. Still University. Retrieved November 27, 2019, from Medical Education website: http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=0f8d38a0-2b46-48be-a145-0d2f3231ca58%40sessionmgr4007

Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *Internet and Higher Education*, *22*, 37–50.

https://doi.org/10.1016/j.iheduc.2014.04.003

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education*, *29*(7), 871–893. https://doi.org/10.1080/09500690600909091

Knowles, M. S., Holton III, E. F., & Swanson, R. A. (2015). *The Adult Learner: The definitive classic in adult education and human resource development* (8th ed.). Routledge.

Korstjens, I., & Moser, A. (2018). European Journal of General Practice Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

Kuhn, T. S. (1996). *The Structure of Scientific Revolutions* (3rd ed.). Chicago: The University of Chicago Press.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education*, *31*(1), 30–43. https://doi.org/10.1080/00220480009596759

Lewin, K. (1936). *Principles of Topological Psychology*. New York, NY: McGraw Hill.

Li, J. C. H. (2016). Effect size measures in a two-independent-samples case with nonnormal and nonhomogeneous data. *Behavior Research Methods*, *48*(4), 1560–1574. https://doi.org/10.3758/s13428-015-0667-z

Limbacher, P. C., & Rosenshine, B. (1972). *Relationship of High-Inference and Low-Inference Observation Measures*.

Limniou, M., Schermbrucker, I., & Lyons, M. (2018). Traditional and flipped classroom approaches delivered by two different teachers: the student perspective. *Education and Information Technologies*, *23*(2), 797–817. https://doi.org/10.1007/s10639-017-9636-8

Lizzio, A. L. F., Wilson, K., & Simons, R. (2002). Studies in Higher Education and

Academic Outcomes : University Students ' Perceptions of the Learning Environment and Academic Outcomes : implications for theory and practice. *Studies in Higher Education*, *27*(1), 27–52. https://doi.org/10.1080/0307507012009935

Logan, K. A., Crump, B. J., Eonie, L. ´, & Rennie, J. (2006). MEASURING THE COMPUTER CLASSROOM ENVIRONMENT: LESSONS LEARNED FROM USING A NEW INSTRUMENT. *Learning Environments Research*, *9*, 67–93. https://doi.org/10.1007/s10984-005-9004-2

Lundin, M., Bergviken Rensfeldt, A., Hillman, T., Lantz-Andersson, A., & Peterson, L. (2018). Higher education dominance and siloed knowledge: a systematic review of flipped classroom research. *International Journal of Educational Technology in Higher Education*, Vol. 15. https://doi.org/10.1186/s41239-018-0101-6

MacFarland, T. W., & Yates, J. M. (2016a). Mann–Whitney U Test. In *Introduction to Nonparametric Statistics for the Biological Sciences Using R* (pp. 103–132). https://doi.org/10.1007/978-3-319-30634-6_4

MacFarland, T. W., & Yates, J. M. (2016b). Wilcoxon Matched-Pairs Signed-Ranks Test. In *Introduction to Nonparametric Statistics for the Biological Sciences Using R* (pp. 133–175). https://doi.org/10.1007/978-3-319-30634-6_5

Martini, F. H., Nath, J. L., & Bartholomew, E. F. (2011). *Fundamentals of Anatomy & Physiology* (9th ed.). Benjamin Cummings.

Mattheis, A., & Jensen, M. (2014). Fostering improved anatomy and physiology instructor pedagogy. *Advances in Physiology Education*, *38*(4), 321–329. https://doi.org/10.1152/advan.00061.2014

Maxwell, J. A. (2013). Conceptual Framework: What do you think is going on? In *Qualitative Research Design: An Interactive Approach* (pp. 39–72). https://doi.org/10.1007/978-3-8349-6169-3_3

McKinley, M., O'Loughlin, V., & Bidle, T. (2019). *Anatomy & Physiology: An Integrative Approach* (3rd ed.). McGraw Hill.

McLean, S., Attardi, S. M., Faden, L., & Goldszmidt, M. (2016). Flipped classrooms and student learning: Not just surface gains. *Advances in Physiology Education*, *40*(1), 47–55. https://doi.org/10.1152/advan.00098.2015

McNally, B., Chipperfield, J., Dorsett, P., Del Fabbro, L., Frommolt, V., Goetz, S., … Rung, A. (2017). Flipped classroom experiences: student preferences and flip strategy in a higher education context. *Higher Education*, *73*(2), 281–298. https://doi.org/10.1007/s10734-016-0014-z

Megaw, P. L., & Zimanyi, M. A. (2019). Redesigning first year anatomy and physiology subjects for allied health students: Introducing active learning experiences for physiology in a first semester subject. *International Journal of Innovation in Science and Mathematics Education*, *27*(8), 26–35.

Michael, J. (2006). Where's the evidence that active learning works? *American Journal of Physiology - Advances in Physiology Education*, *30*(4), 159–167. https://doi.org/10.1152/advan.00053.2006

Miller, S. A., Perrotti, W., Silverthorn, D. U., Dalley, A. F., & Rarey, K. E. (2002). From college to clinic: Reasoning over memorization is key for understanding anatomy. *Anatomical Record*, *269*(2), 69–80. https://doi.org/10.1002/ar.10071

Mishler, E. G., Stern, G. G., Stein, M. I., & Bloom, B. S. (1957). Methods in Personality Assessment. *The Milbank Memorial Fund Quarterly*, Vol. 35, p. 82. https://doi.org/10.2307/3348538

Moos, R. H. (1968). The Assessment of the Social Climates of Correctional Institutions. *Journal of Research in Crime and Delinquency*, *5*(2), 174–188. https://doi.org/10.1177/002242786800500207

Moos, R. H. (1976). *The human context : environmental determinants of behavior : Moos, Rudolf H., 1934- : Free Download, Borrow, and Streaming : Internet Archive*. Retrieved from https://archive.org/details/humancontextenvi00moos

Moos, R. H. (1980). Evaluating classroom learning environments. *Studies in Educational Evaluation*, *6*(3), 239–252. https://doi.org/10.1016/0191-491X(80)90027-9

Morton, D. A., & Colbert-Getz, J. M. (2017). Measuring the impact of the flipped anatomy classroom: The importance of categorizing an assessment by Bloom's taxonomy. *Anatomical Sciences Education*, *10*(2). https://doi.org/10.1002/ase.1635

Murray, H. A., & McAdams, D. P. (2010). Explorations in Personality. *Explorations in Personality*, 1–810. https://doi.org/10.1093/acprof:oso/9780195305067.001.0001

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *Internet and Higher Education*, *25*, 85–95. https://doi.org/10.1016/j.iheduc.2015.02.002

O'Shea, P. M. (2020a). Current issues and future directions. In Z. Walker, D. Tan, & N. K. Koh (Eds.), *Flipped Classrooms with Diverse Learners* (pp. 287–295). https://doi.org/10.4324/9780429506659-5

O'Shea, P. M. (2020b). Flipped learning at university level. In Z. Walker, D. Tan, & N. K. Koh (Eds.), *Flipped Classrooms with Diverse Learners* (pp. 171–181). Springer Singapore.

Older, J. (2004). Anatomy: a must for teaching the next generation. *The Surgeon*, *2*(2), 79–90.

Pandey, P., & Zimitat, C. (2007). Medical students' learning of anatomy: Memorisation, understanding and visualisation. *Medical Education*, *41*(1), 7–14. https://doi.org/10.1111/j.1365-2929.2006.02643.x

Persky, A. M., & McLaughlin, J. E. (2017). The flipped classroom – from theory to practice in health professional education. *American Journal of Pharmaceutical Education*, Vol. 81. https://doi.org/10.5688/ajpe816118

Peter, M., Khoo, E. G. L., Scott, J. B., & Round, W. H. (2016). Learning threshold concepts in an undergraduate engineering flipped classroom. *Deanz2016*, 111–115.

Podolskiy, A. I. (2012). Zone of Proximal Development. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 3485–3487). https://doi.org/10.1007/978-1-4419-1428-6_316

Prashar, A. (2015). Assessing the Flipped Classroom in Operations Management: A Pilot Study. *Journal of Education for Business*, *90*(3), 126–138. https://doi.org/10.1080/08832323.2015.1007904

Priyadharshini, N. A., Kumar, D. V, & Devi, R. (2019). Relevance of Learning Anatomy to Clinical Practice: Perceptive of Medical Students, Interns, and Clinicians. *National Journal of Clinical Anatomy*, *8*(1), 32–37. https://doi.org/10.1055/s-0039-1688893

Quine, W. V. (1948). On What There Is. *The Review of Metaphysics*, Vol. 2, pp. 21–38. https://doi.org/10.2307/20123117

Raftery, A. T. (2007). Anatomy teaching in the UK. *Surgery*, Vol. 25, pp. 1–2. https://doi.org/10.1016/j.mpsur.2006.11.002

Reid, N. (2015). Attitude research in science education. In *Attitude Measurements in Science Education* (pp. 3–46).

Rentoul, A. J., & Fraser, B. J. (1979). Conceptualization of Enquiry-Based or Open Classroom Learning Environments. *Journal of Curriculum Studies*, *11*(3), 233–245. https://doi.org/10.1080/0022027790110306

Richardson, D. R. (2000). Comparison of naive and experienced students of elementary physiology on performance in an advanced course. *American Journal of Physiology - Advances in Physiology Education*, *23*(1), 91–95. https://doi.org/10.1152/advances.2000.23.1.s91

Rotellar, C., & Cain, J. (2016). Research, Perspectives, and Recommendations on Implementing the Flipped Classroom. *American Journal of Pharmaceutical Education, (Article 34)*, *80*(2), 1–9.

Saladin, K. (2018). *Anatomy & Physiology: a unity of form & function 8th edition.* McGraw Hill.

Şengel, E. (2016). To FLIP or not to FLIP: Comparative case study in higher education in Turkey. *Computers in Human Behavior*, *64*, 547–555. https://doi.org/10.1016/j.chb.2016.07.034

Silverman, D. (2006). *Interpreting qualitative data* (3rd ed.). London: SAGE Publications.

Singh, K., Bharatha, A., Sa, B., Adams, O. P., & Majumder, M. A. A. (2019). Teaching anatomy using an active and engaging learning strategy. *BMC Medical Education*, *19*(1). https://doi.org/10.1186/s12909-019-1590-2

Skordi, P., & Fraser, B. J. (2019). Validity and use of the What Is Happening In this Class? (WIHIC) questionnaire in university business statistics classrooms. *Learning Environments Research*, *22*(2), 275–295. https://doi.org/10.1007/s10984-018-09277-4

Slominski, T., Grindberg, S., & Momsen, J. (2019). Physiology is hard: A replication study of students' perceived learning difficulties. *Advances in Physiology Education*, *43*(2), 121–127. https://doi.org/10.1152/advan.00040.2018

Smith, C. F., Martinez-Álvarez, C., & Mchanwell, S. (2014). The context of learning anatomy: Does it make a difference? *Journal of Anatomy*, *224*(3), 270–278. https://doi.org/10.1111/joa.12089

Smith, C. F., & Mathias, H. S. (2011). What impact does anatomy education have on clinical practice? *Clinical Anatomy*, *24*(1), 113–119. https://doi.org/10.1002/ca.21065

Spinner, H., & Fraser, B. J. (2005, June). Evaluation of an innovative mathematics program in terms of classroom environment, student attitudes, and conceptual development. *International Journal of Science and Mathematics Education*, Vol. 3, pp. 267–293. https://doi.org/10.1007/s10763-004-6531-8

Stetzik, L., Deeter, A., Parker, J., & Yukech, C. (2015). Puzzle-based versus traditional lecture: Comparing the effects of pedagogy on academic performance in an undergraduate human anatomy and physiology II lab. *BMC Medical Education*, *15*(1), 1–11. https://doi.org/10.1186/s12909-015-0390-6

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, *15*(2), 171–193. https://doi.org/10.1007/s10984-012-9108-4

Sturges, D., & Maurer, T. (2013). Allied Health Students' Perceptions of Class Difficulty: The Case of Undergraduate Human Anatomy and Physiology Classes. *The Internet Journal of Allied Health Sciences and Practice*, *11*(4).

Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: A review for its modernization. *Anatomical Sciences Education*, *3*(2), 83–93. https://doi.org/10.1002/ase.139

Sweller, J., Ayres, P., Kalyuga, S., Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. In *Cognitive Load Theory*. https://doi.org/10.1007/978-1-4419-8126-4_2

Taylor, P. C., Fraser, B. J., & Fisher, D. L. (1997). Monitoring constructivist classroom learning environments. *International Journal of Educational Research*, *27*(4), 293–302. https://doi.org/10.1016/S0883-0355(97)90011-2

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. In *TRENDS in Sport Sciences* (Vol. 1). Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1068.5305&rep=rep1&type=pdf

Tytler, R., & Osborne, J. (2012). Student attitudes and aspirations towards science. In *Second International Handbook of Science Education* (pp. 597–625). https://doi.org/10.1007/978-1-4020-9041-7_41

van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, *28*(May), 1–18. https://doi.org/10.1016/j.edurev.2019.05.003

Villafañe, S. M., & Lewis, J. E. (2016). Exploring a measure of science attitude for different groups of students enrolled in introductory college chemistry.

Chemistry Education Research and Practice, *17*(4), 731–742. https://doi.org/10.1039/c5rp00185d

Vitali, J., Blackmore, C., Mortazavi, S., & Anderton, R. (2020). Tertiary Anatomy and Physiology, A Barrier for Student Success. *International Journal of Higher Education*, *9*(2). https://doi.org/10.5430/ijhe.v9n2p289

Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology*, *59*(6 PART 1), 414–419. https://doi.org/10.1037/h0026490

Walker, Z., Tan, D., Klimplová, L., & Bicen, H. (2020). An Introduction to Flipping the Classroom. In *Flipped Classrooms with Diverse Learners* (pp. 3–15). Retrieved from http://web.a.ebscohost.com.dbgw.lis.curtin.edu.au/ehost/ebookviewer/ebook/bmxlYmtfXzI1NDU2NDJfX0FO0?sid=fc1b3f6d-26be-4980-b640-30fb6bc67dc8@sessionmgr4007&vid=8&format=EK&rid=1

Waxer, M., & Morton, J. B. (2012). Cognitive Conflict and Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 585–587). https://doi.org/10.1007/978-1-4419-1428-6_280

Wilhelmsson, N., Dahlgren, L. O., Hult, H., & Josephson, A. (2011). On the anatomy of understanding. *Studies in Higher Education*, *36*(2), 153–165. https://doi.org/10.1080/03075070903514054

Wilhelmsson, N., Dahlgren, L. O., Hult, H., Scheja, M., Lonka, K., & Josephson, A. (2010). The anatomy of learning anatomy. *Advances in Health Sciences Education*, *15*(2), 153–165. https://doi.org/10.1007/s10459-009-9171-5

Woods, N. N., Brooks, L. R., & Norman, G. R. (2007). It all make sense: Biomedical knowledge, causal connections and memory in the novice diagnostician. *Advances in Health Sciences Education*, *12*(4), 405–415. https://doi.org/10.1007/s10459-006-9055-x

Yang, F.-Y., & Tsai, C.-C. (2012). Personal Epistemology and Science Learning: A Review on Empirical Studies. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie

(Eds.), *Second International Handbook of Science Education* (pp. 259–287). Springer.

Young, G. (2011a). Piaget and Development. In *Development and Causality: Neo-Piagetian Perspectives* (pp. 47–72). https://doi.org/10.1007/978-1-4419-9422-6_3

Young, G. (2011b). Piaget and Development. In *Development and Causality: Neo-Piagetian Perspectives* (pp. 47–72). https://doi.org/10.1007/978-1-4419-9422-6_3

Zainuddin, Z., & Perera, C. J. (2019). Exploring students' competence, autonomy and relatedness in the flipped classroom pedagogical model. *Journal of Further and Higher Education*, *43*(1), 115–126. https://doi.org/10.1080/0309877X.2017.1356916

Appendices

Appendix A English Version of Test of Science-Related Attitudes (TOSRA)

NOTE: The TOSRA was developed by Fraser (1978) and is discussed in Sections 2.5.5.3, 3.6.3, and 4.4.2. It was used in my study and included in this thesis with permission of the author.

TOSRA

TEST OF SCIENCE-RELATED ATTITUDES

Barry J. Fraser

DIRECTIONS

1 This test contains a number of statements about science. You will be asked what you yourself think about these statements. There are no 'right' or 'wrong' answers. Your opinion is what is wanted.

2 All answers should be given on the separate Answer Sheet. Please do not write on this booklet.

3 For each statement, draw a circle around

- SA if you **STRONGLY AGREE** with the statement;
- A if you **AGREE** with the statement;
- N if you are **NOT SURE**;
- D if you **DISAGREE** with the statement;
- SD if you **STRONGLY DISAGREE** with the statement.

Practice Item

0 It would be interesting to learn about boats.

Suppose that you AGREE with this statement, then you would circle A on your Answer Sheet, like this:

0 SA (A) N D SD

4 If you change your mind about an answer, cross it out and circle another one.

5 Although some statements in this test are fairly similar to other statements, you are asked to indicate your opinion about all statements.

Published by
The Australian Council for Educational Research Limited
Radford House, Frederick Street, Hawthorn, Victoria 3122

Typesetting direct from Wang diskette by
Publication Perspectives
200 Cheltenham Road, Dandenong, Victoria 3175

Printed by Allanby Press
1A Crescent Road, Camberwell, Victoria 3124

ACER Australian Council for Educational Research

Curtin University

Page 2

1 Money spent on science is well worth spending.

2 Scientists usually like to go to their laboratories when they have a day off.

3 I would prefer to find out why something happens by doing an experiment than by being told.

4 I enjoy reading about things which disagree with my previous ideas.

5 Science lessons are fun.

6 I would like to belong to a science club.

7 I would dislike being a scientist after I leave school.

8 Science is man's worst enemy.

9 Scientists are about as fit and healthy as other people.

10 Doing experiments is not as good as finding out information from teachers.

11 I dislike repeating experiments to check that I get the same results.

12 I dislike science lessons.

13 I get bored when watching science programs on TV at home.

14 When I leave school, I would like to work with people who make discoveries in science.

15 Public money spent on science in the last few years has been used wisely.

16 Scientists do not have enough time to spend with their families.

17 I would prefer to do experiments than to read about them.

18 I am curious about the world in which we live.

19 School should have more science lessons each week.

20 I would like to be given a science book or a piece of scientific equipment as a present.

21 I would dislike a job in a science laboratory after I leave school.

22 Scientific discoveries are doing more harm than good.

23 Scientists like sport as much as other people do.

24 I would rather agree with other people than do an experiment to find out for myself.

25 Finding out about new things is unimportant.

26 Science lessons bore me.

27 I dislike reading books about science during my holidays.

28 Working in a science laboratory would be an interesting way to earn a living.

Page 3

29 The government should spend more money on scientific research.

30 Scientists are less friendly than other people.

31 I would prefer to do my own experiments than to find out information from a teacher.

32 I like to listen to people whose opinions are different from mine.

33 Science is one of the most interesting school subjects.

34 I would like to do science experiments at home.

35 A career in science would be dull and boring.

36 Too many laboratories are being built at the expense of the rest of education.

37 Scientists can have a normal family life.

38 I would rather find out about things by asking an expert than by doing an experiment.

39 I find it boring to hear about new ideas.

40 Science lessons are a waste of time.

41 Talking to friends about science after school would be boring.

42 I would like to teach science when I leave school.

43 Science helps to make life better.

44 Scientists do not care about their working conditions.

45 I would rather solve a problem by doing an experiment than be told the answer.

46 In science experiments, I like to use new methods which I have not used before.

47 I really enjoy going to science lessons.

48 I would enjoy having a job in a science laboratory during my school holidays.

49 A job as a scientist would be boring.

50 This country is spending too much money on science.

51 Scientists are just as interested in art and music as other people are.

52 It is better to ask the teacher the answer than to find it out by doing experiments.

53 I am unwilling to change my ideas when evidence shows that the ideas are poor.

54 The material covered in science lessons is uninteresting.

55 Listening to talk about science on the radio would be boring.

56 A job as a scientist would be interesting.

57 Science can help to make the world a better place in the future.

58 Few scientists are happily married.

59 I would prefer to do an experiment on a topic than to read about it in science magazines.

60 In science experiments, I report unexpected results as well as expected ones.

61 I look forward to science lessons.

62 I would enjoy visiting a science museum at the weekend.

63 I would dislike becoming a scientist because it needs too much education.

64 Money used on scientific projects is wasted.

65 If you met a scientist, he would probably look like anyone else you might meet.

66 It is better to be told scientific facts than to find them out from experiments.

67 I dislike listening to other people's opinions.

68 I would enjoy school more if there were no science lessons.

69 I dislike reading newspaper articles about science.

70 I would like to be a scientist when I leave school.

Test of Science-Related Attitudes

Answer Sheet

Name ______________________________

School ______________________ Year/Class __________

Page 2	STRONGLY AGREE	AGREE	NOT SURE	DISAGREE	STRONGLY DISAGREE	Page 3	STRONGLY AGREE	AGREE	NOT SURE	DISAGREE	STRONGLY DISAGREE	Page 4	STRONGLY AGREE	AGREE	NOT SURE	DISAGREE	STRONGLY DISAGREE
1	SA	A	N	D	SD	29	SA	A	N	D	SD	50	SA	A	N	D	SD
2	SA	A	N	D	SD	30	SA	A	N	D	SD	51	SA	A	N	D	SD
3	SA	A	N	D	SD	31	SA	A	N	D	SD	52	SA	A	N	D	SD
4	SA	A	N	D	SD	32	SA	A	N	D	SD	53	SA	A	N	D	SD
5	SA	A	N	D	SD	33	SA	A	N	D	SD	54	SA	A	N	D	SD
6	SA	A	N	D	SD	34	SA	A	N	D	SD	55	SA	A	N	D	SD
7	SA	A	N	D	SD	35	SA	A	N	D	SD	56	SA	A	N	D	SD
8	SA	A	N	D	SD	36	SA	A	N	D	SD	57	SA	A	N	D	SD
9	SA	A	N	D	SD	37	SA	A	N	D	SD	58	SA	A	N	D	SD
10	SA	A	N	D	SD	38	SA	A	N	D	SD	59	SA	A	N	D	SD
11	SA	A	N	D	SD	39	SA	A	N	D	SD	60	SA	A	N	D	SD
12	SA	A	N	D	SD	40	SA	A	N	D	SD	61	SA	A	N	D	SD
13	SA	A	N	D	SD	41	SA	A	N	D	SD	62	SA	A	N	D	SD
14	SA	A	N	D	SD	42	SA	A	N	D	SD	63	SA	A	N	D	SD
15	SA	A	N	D	SD	43	SA	A	N	D	SD	64	SA	A	N	D	SD
16	SA	A	N	D	SD	44	SA	A	N	D	SD	65	SA	A	N	D	SD
17	SA	A	N	D	SD	45	SA	A	N	D	SD	66	SA	A	N	D	SD
18	SA	A	N	D	SD	46	SA	A	N	D	SD	67	SA	A	N	D	SD
19	SA	A	N	D	SD	47	SA	A	N	D	SD	68	SA	A	N	D	SD
20	SA	A	N	D	SD	48	SA	A	N	D	SD	69	SA	A	N	D	SD
21	SA	A	N	D	SD	49	SA	A	N	D	SD	70	SA	A	N	D	SD
22	SA	A	N	D	SD												
23	SA	A	N	D	SD												
24	SA	A	N	D	SD												
25	SA	A	N	D	SD												
26	SA	A	N	D	SD												
27	SA	A	N	D	SD												
28	SA	A	N	D	SD												

For Teacher Use Only

S ___ N ___ I ___ A ___ E ___ L ___ C ___

Appendix B College and University Classroom Environment Inventory (CUCEI)

NOTE: The CUCEI was developed by Fraser, Treagust and Dennis (1986) and is discussed in Sections 2.5.3.4, 2.5.4, 3.6.2, and 4.4.1. It was used in my study and included in this thesis with permission of the authors.

Curtin University

College and University Classroom Environment Inventory

The purpose of this questionnaire is to find out your opinions about the class you are attending right now. This questionnaire is designed for use in gathering opinions about classes

This form of the questionnaire assesses your opinion about what this class is actually like. Indicate your opinion about each questionnaire statement by selecting one of the following options:

For research purposes, please indicate:

Your academic major: ________________ Today's date: ___________________________

Year at this college: _________________ Subject of this classroom: __________________

I am male/female: ___________________ My instructor is male/female: ________________

Attitude toward students	1.	The teacher considers students' feelings.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	2.	The teacher talks rather than listens.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	3.	The class is made up of individuals who don't know each other well.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	4.	The students look forward to coming to classes.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	5.	Students know exactly what has to be done in our class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	6.	New ideas are seldom tried out in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	7.	All students in the class are expected to do the same work in the same way and in the same time.	Strongly Agree	Agree	Disagree	Strongly Disagree
Attitude toward students	8.	The teacher talks individually with students.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	9.	Students put effort into what they do in class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	10.	Each student knows the other members of the class by their first names.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	11.	Students are dissatisfied with what is done in the class,	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	12.	Getting a certain amount of work done is important in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	13.	New and different ways of teaching are seldom used in own pace this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	14.	Students are generally allowed to	Strongly	Agree	Disagree	Strongly

		work at their own pace.	Agree			Disagree
Attitude toward students	15.	The teacher goes out of his/her way to help students. in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	16.	Students 'clockwatch' in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	17.	Friendships are made among students in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	18.	After the class the students have a sense of satisfaction.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	19.	The group often gets sidetracked instead of sticking to the point.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	20.	The teacher thinks up innovative activities for students to do.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	21.	Students have a say in how class time is spent.	Strongly Agree	Agree	Disagree	Strongly Disagree
Attitude toward students	22.	The teacher helps each student who is having trouble.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	23.	Students in this class pay attention to what others are saying.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	24.	Students don't have much chance to get to know each other in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	25.	Classes are a waste of time.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	26.	This is a disorganized class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	27.	Teaching approaches in this class are characterized by innovation and variety.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	28.	Students are allowed to choose activities and how they will work.	Strongly Agree	Agree	Disagree	Strongly Disagree
Attitude toward students	29.	The teacher seldom moves around the classroom to talk with new students.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	30.	Students seldom present their work to the class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	31.	It takes a long time to get to know everybody by his/her first name in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	32.	Classes are boring.	Strongly Agree	Agree	Disagree	Strongly Disagree

Class organization	33.	Class assignments are clear so everyone knows what to do.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	34.	The seating in this class is arranged in the same way each week.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	35.	Teaching approaches allow students to proceed at their own pace.	Strongly Agree	Agree	Disagree	Strongly Disagree
Attitude toward students	36.	The teacher isn't interested in students' problems.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	37.	There are opportunities for students to express opinions.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	38.	Students in this class get to know each other well.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	39.	Students enjoy going to this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	40.	This class seldom starts on time.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	41.	The teacher often thinks of unusual class activities.	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	42.	There is little opportunity for a student to pursue his/her particular interest in this class.	Strongly Agree	Agree	Disagree	Strongly Disagree
Attitude toward students	43.	The teacher is unfriendly and inconsiderate towards students	Strongly Agree	Agree	Disagree	Strongly Disagree
Autonomy-power sharing	44.	The teacher dominates class discussion.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student-student relationships	45.	Students in this class aren't very interested in getting to know other students.	Strongly Agree	Agree	Disagree	Strongly Disagree
Student interest-motivation	46.	Classes are interesting.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	47.	Activities in this class are clearly and carefully planned.	Strongly Agree	Agree	Disagree	Strongly Disagree
Class organization	48.	Students seem to do the same type of activities every class.	Strongly Agree	Agree	Disagree	Strongly Disagree

Appendix C Informed Consent Form

Curtin University

Curtin University is a trademark of Curtin University of Technology.
CRICOS Provider Code 00301J (WA), 02637B (NSW)

Faculty of Humanities. School of Education
GPO BOX U1987 Perth, WA 6845

Letter of Consent

A Flipped Classroom Approach in Undergraduate Anatomy and Physiology: A Mixed Methods Study Evaluating Learning Environment and Student Outcomes

I hereby give my consent to Scott Gaines M.S., a researcher/student in the Faculty of Humanities, School of Education at Curtin University to record and document my participation activities.

I therefore give permission for the use of this data, and other information which I have agreed may be obtained or requested, in the writing up of this study, subject to the following conditions:

....

....

....

....

....

I understand that this project has been approved by Curtin University Human Research Ethics Committee and will be carried out in line with the National Statement on Ethical Conduct in Human Research (2007). I understand I will receive a copy of this Information Statement and Consent Form.

My participation in this study is voluntary, and I understand that I may withdraw from the study at any time.

SIGNATURES

Participant .... Date ....

Researcher .... Date ....

Appendix D Introduction Letter

Curtin University

Curtin University is a trademark of Curtin University of Technology.
CRICOS Provider Code 00301J (WA), 02637B (NSW)

Faculty of Humanities. School of Education
GPO BOX U1987 Perth, WA 6845

Letter of Introduction

7th of January, 2019

Dear CUI BIO 246/247 Student,

This letter is to introduce Scott Gaines, M.S. who is a Science Education student in the School of Education at Curtin University. I will produce my current student card, as proof of identity.
I am undertaking research leading to the production of a thesis or other publications on the subject of A Flipped Classroom Approach in Undergraduate Anatomy and Physiology: A Mixed Methods Study Evaluating Learning Environment and Student Outcomes. The goal of this research is to determine the effect of a flipped classroom model in a content-dense undergraduate science course such as anatomy and physiology.

Therefore, it was be greatly appreciated if you would volunteer to assist in this project by participating completing two brief questionnaires (approximately 30 min., given at the end of lab) and participating in a brief (30 min.) focus group after completion of the term (in the Fall 2019).

Please also be assured that all information provided will be treated with the strictest confidence and no participants will be individually identified in the resulting thesis/publication. You are, of course entirely fee to discontinue from this project at any time or to decline to answer particular questions.

PLEASE NOTE: This is participation is completely voluntary and anonymous (Prof. Gaines will NOT know who participates and who does not). Your participation or non-participation has no effect on your grade. Both questionnaires and the focus groups will be administered by researchers other than Prof. Gaines.

As it is my intention to record the focus group, an additional consent form is attached that permits the recording of the interview and to use the recording or a transcription in the preparation of the thesis/publication on the condition that your name or identity is not revealed (please note: Prof. Gaines will only receive the transcripts, not the original recordings). Please be aware however, that it may be necessary to make the recording available to other researchers on the same conditions as well as to secretarial assistants for transcription purposes only. Under these circumstances, such persons will be conditioned to the same confidentiality obligations.

Any enquiries you may have concerning this project should be directed to me at
scott.gaines@postgrad.curtin.edu.au or scott.gaines@cui.edu

Thank you for your attention and assistance

Yours sincerely

Scott Gaines, M.S.
Resident Faculty,
Concordia University Irvine

P. John Williams, Ph.D.
Supervisor
Director of Graduate Studies in the School of Education at Curtin University

Curtin University Human Research Ethics Committee (HREC) has approved this study (HREC number 2018-0468). Should you wish to discuss the study with someone not directly involved, in particular, any matters concerning the conduct of the study or your rights as a participant, or you wish to make a confidential complaint, you may contact the Ethics Officer on (08) 9266 9223 or the Manager, Research Integrity on (08) 9266 7093 or email hrec@curtin.edu.au.

Appendix E College and University Classroom Environment Inventory (CUCEI) data

Average Initial (Pre-) CUCEI Scores for Flipped Classroom

	CUCEI Scale						
Student	Personali-sation	Involvement	Student Cohesiveness	Satisfaction	Task Orientation	Innovation	Individuali-zation
1	4.86	4.71	4.57	4.86	4.86	4.57	2.86
2	4.57	3.43	3.29	4.29	4.29	3.00	2.43
3	4.71	3.57	4.71	4.00	3.57	3.29	3.14
4	4.14	3.29	4.00	4.00	4.14	3.29	3.14
5	4.86	4.00	3.43	3.57	4.57	3.14	3.29
6	4.00	3.14	3.29	3.71	3.43	2.57	2.71
7	4.29	3.43	3.43	4.14	4.43	3.43	2.71
8	4.86	4.71	4.57	4.57	5.00	3.43	2.57
9	N/A	N/A	N/A	N/A	N/A	N/A	N/A
10	4.00	3.14	4.00	3.71	4.14	3.14	3.57
11	4.29	3.29	2.86	3.43	3.86	3.29	2.57
12	4.00	3.86	3.57	4.71	4.00	3.00	3.29
13	4.86	4.71	4.57	4.86	4.86	4.57	2.86
Mean	4.45	3.77	3.86	4.15	4.26	3.39	2.93
SD	0.37	0.62	0.63	0.50	0.51	0.60	0.35

Average Final (Post-) CUCEI Scores for Flipped Classroom

	CUCEI Scale						
Student	Personali-sation	Involvement	Student Cohesiveness	Satisfaction	Task Orientation	Innovation	Individuali-zation
1	4.43	3.57	3.00	4.14	3.86	2.57	3.29
2	4.71	3.86	3.86	4.00	4.29	3.14	3.00
3	4.71	3.14	4.57	3.57	3.43	2.71	3.14
4	4.57	3.86	4.29	3.57	4.14	3.29	3.00
5	5.00	3.57	3.14	4.14	4.14	3.14	3.00
6	3.29	3.00	3.00	4.00	3.57	2.29	3.43
7	4.43	3.43	3.43	4.00	3.86	3.57	2.71
8	4.86	4.43	4.43	4.86	5.00	3.71	2.43
9	N/A	N/A	N/A	N/A	N/A	N/A	N/A
10	4.29	3.00	3.14	3.00	3.86	2.71	3.14
11	3.86	3.43	2.86	4.14	3.57	3.71	3.00
12	4.29	4.14	4.14	4.57	4.29	3.14	3.86
13	5.00	4.43	4.71	4.86	4.71	4.29	3.00
Mean	4.45	3.65	3.71	4.07	4.06	3.19	3.08
SD	0.49	0.50	0.69	0.54	0.47	0.57	0.35

Average Initial (Pre-) CUCEI Scores for Traditional Classroom

	CUCEI Scale						
Student	Personali-sation	Involvement	Student Cohesiveness	Satisfaction	Task Orientation	Innovation	Individuali-zation
1	4.57	3.29	2.29	4.29	4.29	2.86	2.71
2	4.00	2.86	3.71	4.00	4.00	3.14	2.71
3	4.57	2.57	4.00	2.86	2.71	2.71	2.71
4	4.57	3.71	4.71	4.57	4.71	3.00	2.43
5	4.14	3.43	3.14	3.43	3.71	2.29	3.00
6	4.43	3.86	3.71	4.43	4.43	3.00	3.00
7	4.57	3.43	3.86	4.43	4.29	3.43	2.43
8	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9	4.14	2.57	3.14	3.71	4.00	2.29	3.29
10	2.86	3.14	2.00	4.00	4.29	3.14	2.00
11	4.43	2.29	3.57	3.57	4.14	2.57	2.00
12	3.71	4.14	3.43	3.71	4.00	2.86	3.14
13	4.00	3.14	3.43	2.57	3.43	3.43	3.00
Mean	4.17	3.20	3.42	3.80	4.00	2.89	2.70
SD	0.50	0.56	0.73	0.63	0.52	0.38	0.42

Average Final (Post-) CUCEI Scores for Traditional Classroom

	CUCEI Scale						
Student	Personali-sation	Involvement	Student Cohesiveness	Satisfaction	Task Orientation	Innovation	Individuali-zation
1	4.29	2.43	2.00	2.86	2.86	2.29	2.43
2	4.00	2.86	3.71	3.71	4.00	2.86	2.71
3	4.71	3.86	4.86	4.29	4.14	3.14	3.00
4	4.71	3.14	3.29	4.43	4.14	2.86	2.43
5	3.86	3.14	3.43	3.00	3.29	2.29	3.14
6	N/A	N/A	N/A	N/A	N/A	N/A	N/A
7	4.14	2.86	3.43	4.14	4.29	3.29	2.14
8	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9	4.14	2.71	2.86	2.29	3.86	1.57	2.43
10	3.71	2.86	2.29	4.00	3.57	2.00	2.43
11	4.71	3.00	3.14	3.71	4.71	2.14	2.29
12	4.14	4.00	3.86	3.86	3.29	2.71	3.43
13	4.00	3.14	3.43	2.57	3.43	3.43	3.00
Mean	4.22	3.09	3.30	3.53	3.78	2.60	2.68
SD	0.35	0.47	0.77	0.73	0.54	0.58	0.41

Appendix F Test of Science-Related Attitudes (TOSRA) data

Average Initial (Pre-) TOSRA Scores for Flipped Classroom

Student	Social Implications of Science	Normality of Scientists	Attitude to Scientific Inquiry	Adoption of Scientific Attitudes	Enjoyment of Science Lessons	Leisure Interest in Science	Career Interest in Science
1	4.7	3.4	3.9	4.1	3.8	3.6	3.8
2	4	3	2.4	3.7	4.3	2.9	3.9
3	4.3	4.3	3.4	3.8	4	2.1	2.9
4	4.3	4	3.6	4	5	4.6	4.9
5	4.7	4.7	3.2	3.8	4.8	4.2	3.9
6	3.8	3.2	3.1	3.4	2.7	2.1	2.5
7	3.8	3.9	4.1	4.1	3.9	3.3	3.9
8	4.5	3.8	4.9	4.7	4.9	3.3	3.6
9	N/A	N/A	N/A	N/A	N/A	N/A	N/A
10	3.9	3.4	2.9	3.5	4.2	2.9	3.5
11	3.4	3	2.4	3.3	3	1.3	1.5
12	3.9	2.9	3.5	4.1	4.8	4.5	3.6
13	4.6	4.3	4.1	4.7	4.7	4.3	3.9
Mean	4.16	3.66	3.46	3.93	4.18	3.26	3.49
SD	0.42	0.59	0.73	0.45	0.74	1.05	0.86

Average Initial (Pre-) TOSRA Scores for Flipped Classroom

Student	Social Implications of Science	Normality of Scientists	Attitude to Scientific Inquiry	Adoption of Scientific Attitudes	Enjoyment of Science Lessons	Leisure Interest in Science	Career Interest in Science
1	4.6	4.2	4.2	3.9	4.7	4.3	4.3
2	4.2	3.2	3.1	3.5	3.5	2.9	3
3	4.4	4	2.8	3.6	3.2	1.9	2.9
4	4.1	3.6	4.4	4.1	4.3	4.1	4.4
5	4.9	4.6	3.6	4.5	4.4	4	3.8
6	3.2	3.2	3.3	3.2	3.3	3.2	2.7
7	4.5	4.2	4	4.7	4.2	3.9	4.5
8	4.5	4.4	4.7	4.4	4.9	3.3	3.9
9	N/A	N/A	N/A	N/A	N/A	N/A	N/A
10	4	3.2	2.4	4	4.3	2.7	3.4
11	3.5	3	2	3.6	3.6	2.8	2.8
12	4.2	3.4	3.1	4.2	4.9	4.2	3.7
13	4.5	4.2	3.8	4.5	4.8	4.2	3.7
Mean	4.22	3.77	3.45	4.02	4.18	3.46	3.59
SD	0.48	0.56	0.82	0.47	0.63	0.77	0.63

Average Initial (Pre-) TOSRA Scores for Traditional Classroom

Student	Social Implications of Science	Normality of Scientists	Attitude to Scientific Inquiry	Adoption of Scientific Attitudes	Enjoyment of Science Lessons	Leisure Interest in Science	Career Interest in Science
1	4.9	3.3	3	3.8	5	4.4	4.9
2	3.7	3.8	3.3	4.1	3.7	3.1	2.9
3	3.9	3	4.5	4.5	4	3.1	2.2
4	3.9	3.8	4.3	4.2	4.8	3.6	3.3
5	4.2	3.9	3.9	3.6	3.8	2.9	3.8
6	3.7	3.6	3.5	4	3.6	3.1	4
7	3.6	4.2	2.4	3.9	4.4	3.3	3.9
8	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9	3.9	3.2	3	3.5	4.6	3.9	3.1
10	3.8	3.6	3.9	3	3.9	3.3	3.5
11	4.3	4.2	3	4.4	4.3	3.8	3.6
12	4	3.4	4.4	3.9	4	3.1	3.5
13	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mean	3.99	3.64	3.56	3.90	4.19	3.42	3.52
SD	0.37	0.39	0.69	0.43	0.46	0.46	0.69

Average Initial (Post-) TOSRA Scores for Traditional Classroom

Student	Social Implications of Science	Normality of Scientists	Attitude to Scientific Inquiry	Adoption of Scientific Attitudes	Enjoyment of Science Lessons	Leisure Interest in Science	Career Interest in Science
1	4.5	3.3	3.4	3.5	4.9	4.5	4.7
2	3.8	3.7	3.5	3.9	3.7	3.4	2.9
3	3.9	3.3	5	4.1	4.9	3.3	2.7
4	3.9	4	4	4.2	4.6	3.2	3
5	3.7	2.5	3.6	3.5	3.4	2.6	3.4
6	N/A	N/A	N/A	N/A	N/A	N/A	N/A
7	4.4	4.4	1.9	3.5	4.7	3.3	3.8
8	N/A	N/A	N/A	N/A	N/A	N/A	N/A
9	4.1	3.2	3.4	4.2	4.2	4	3.7
10	3.6	3.5	3.1	3	3.8	3.2	3.6
11	4.1	4.3	4.1	3.9	4	3.7	3.3
12	4.2	3.1	3.9	4	3.8	3.6	3.3
13	4	3	3.6	4.6	4	4	3.6
Mean	4.02	3.53	3.59	3.78	4.20	3.48	3.44
SD	0.29	0.58	0.80	0.39	0.54	0.51	0.57

Appendix G Examination and Laboratory Practical Data

Examination Scores for Traditional Classroom

Student	Examination 1 Endocrine System	Examination 2 Digestive System	Examination 3 Blood and Heart	Examination 4 Vessels, Lymph and Immune	Examination 5 Respiratory, Renal, Fluids	Final Examination
A	61.53	57.10	54.39	70.07	54.82	47.01
B	93.21	88.82	88.32	90.51	83.22	90.61
C	89.55	83.00	86.10	83.71	65.20	77.54
D	75.27	80.97	82.03	76.29	72.23	70.79
E	71.64	72.75	76.11	85.19	70.84	73.97
F	76.92	77.46	88.50	91.49	75.78	79.46
G	89.58	85.53	95.73	95.79	86.68	88.89
H	89.14	92.53	78.17	77.21	73.68	74.63
I	91.70	95.92	81.73	90.47	86.06	92.40
J	68.98	61.87	90.80	93.87	58.68	70.92
K	95.19	89.35	93.97	98.08	83.58	92.49
L	52.14	70.86	56.05	69.78	N/A	N/A
M	73.00	75.18	62.53	70.18	59.68	59.93
Average %	79.06	79.33	79.57	84.05	72.54	76.55
Standard Deviation	13.50	11.65	13.83	10.29	11.11	13.78

Examination Scores for Flipped Classroom

Student	Examination 1 Endocrine System	Examination 2 Digestive System	Examination 3 Blood and Heart	Examination 4 Vessels, Lymph and Immune	Examination 5 Respiratory, Renal, Fluids	Final Examination
A	88.73	71.33	73.72	68.97	58.28	46.51
B	85.33	79.10	88.01	93.93	69.88	78.57
C	65.04	38.17	54.08	54.62	45.88	51.00
D	92.01	96.04	79.82	86.09	80.08	87.62
E	88.06	90.05	87.53	80.54	80.14	82.94
F	74.43	77.77	94.40	97.80	69.32	68.11
G	80.41	76.25	76.01	84.53	61.03	79.17
H	93.70	92.01	88.65	87.82	68.57	96.58
I	68.56	73.20	63.37	78.08	59.01	73.88
J	72.27	75.85	68.25	68.72	77.04	79.87
K	50.03	58.77	83.25	87.12	65.34	57.93
L	66.76	75.53	73.75	77.93	66.37	74.82
M	61.45	79.44	57.21	67.47	55.51	65.50
Average %	75.91	75.65	76.00	79.51	65.88	72.50
Standard Deviation	13.38	14.82	12.60	12.03	10.00	14.37

Practical Examination Scores for Traditional Classroom

Student	PRACTICAL EXAMINATION #1	PRACTICAL EXAMINATION #2	PRACTICAL EXAMINATION #3
A	30.50	30.00	17.00
B	53.50	80.00	31.50
C	56.50	85.50	31.00
D	44.50	59.50	30.50
E	42.00	67.00	29.00
F	49.00	62.50	32.00
G	61.50	76.50	36.00
H	58.50	81.00	31.00
I	65.00	90.50	36.00
J	36.50	51.50	28.00
K	49.50	73.50	37.00
L	40.00	N/A	N/A
M	47.00	73.00	24.50
Average Score	48.77	69.21	30.29
Standard Deviation	10.13	16.67	5.50

Practical Examination Scores for Flipped Classroom

Student	PRACTICAL EXAMINATION #1	PRACTICAL EXAMINATION #2	PRACTICAL EXAMINATION #3
A	47.50	55.00	19.00
B	45.00	67.00	34.00
C	22.50	19.00	23.50
D	54.50	84.00	22.50
E	48.50	74.50	36.50
F	40.00	23.00	19.50
G	51.50	71.50	32.00
H	52.50	80.00	37.00
I	32.50	52.00	29.00
J	44.50	68.00	33.00
K	45.00	63.50	24.50
L	51.00	52.50	29.00
M	31.50	44.50	18.00
Average Score	43.58	58.04	27.50
Standard Deviation	9.52	20.06	6.74

Appendix H Interview Questions

Semi-Structured Focus Group Questions

- Overall, how do you feel regarding the flipped classroom approach?
- Compared to the first semester of Anatomy and Physiology, taught in the traditional format, which structure do you feel worked best for you?
- Do you have a preference of types of video lectures (created by the instructor or found on YouTube) used in the course?
- Do you feel the activities in class assisted in your learning of Anatomy and Physiology?
- How helpful do you feel the case studies were in learning Anatomy and Physiology?
- How helpful do you feel the POGILs were in learning Anatomy and Physiology?
- How important do you feel interacting with fellow students is to the learning process?
- What would be your recommendations (if any) for the instructor to make the class better in the future?

Appendix I Curtin Human Research Ethics Approval

Curtin University

Office of Research and Development

GPO Box U1987
Perth Western Australia 6845

Telephone +61 8 9266 7863
Facsimile +61 8 9266 3793
Web research.curtin.edu.au

19-Jul-2018

Name: John Williams
Department/School: School of Education
Email: Pjohn.Williams@curtin.edu.au

Dear John Williams

RE: Ethics Office approval
Approval number: HRE2018-0468

Thank you for submitting your application to the Human Research Ethics Office for the project **A Flipped Classroom Approach in Undergraduate Anatomy and Physiology: A Mixed Methods Study Evaluating Classroom Environment and Student Attitudes**.

Your application was reviewed through the Curtin University Low risk review process.

The review outcome is: **Approved**.

Your proposal meets the requirements described in the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research (2007)*.

Approval is granted for a period of one year from **19-Jul-2018** to **18-Jul-2019**. Continuation of approval will be granted on an annual basis following submission of an annual report.

Personnel authorised to work on this project:

Name	Role
Gaines, Scott	Student
Williams, John	CI
Fraser, Barry	Supervisor

Approved documents:

Document

Standard conditions of approval

1. Research must be conducted according to the approved proposal
2. Report in a timely manner anything that might warrant review of ethical approval of the project including:

- proposed changes to the approved proposal or conduct of the study
- unanticipated problems that might affect continued ethical acceptability of the project
- major deviations from the approved proposal and/or regulatory guidelines
- serious adverse events

3. Amendments to the proposal must be approved by the Human Research Ethics Office before they are implemented (except where an amendment is undertaken to eliminate an immediate risk to participants)
4. An annual progress report must be submitted to the Human Research Ethics Office on or before the anniversary of approval and a completion report submitted on completion of the project
5. Personnel working on this project must be adequately qualified by education, training and experience for their role, or supervised
6. Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, that bears on this project
7. Changes to personnel working on this project must be reported to the Human Research Ethics Office
8. Data and primary materials must be retained and stored in accordance with the Western Australian University Sector Disposal Authority (WAUSDA) and the Curtin University Research Data and Primary Materials policy
9. Where practicable, results of the research should be made available to the research participants in a timely and clear manner
10. Unless prohibited by contractual obligations, results of the research should be disseminated in a manner that will allow public scrutiny; the Human Research Ethics Office must be informed of any constraints on publication
11. Approval is dependent upon ongoing compliance of the research with the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research, applicable legal requirements, and with Curtin University policies, procedures and governance requirements
12. The Human Research Ethics Office may conduct audits on a portion of approved projects.

Special Conditions of Approval

None

This letter constitutes low risk/negligible risk approval only. This project may not proceed until you have met all of the Curtin University research governance requirements.

Should you have any queries regarding consideration of your project, please contact the Ethics Support Officer for your faculty or the Ethics Office at hrec@curtin.edu.au or on 9266 2784.

Yours sincerely

Catherine Gangell
Manager, Research Integrity

Appendix J Concordia University Irvine IRB Approval

CONCORDIA
UNIVERSITY IRVINE

INSTITUTIONAL REVIEW BOARD DECISION

☐ Exempt Review 45 CFR 46.101 ☑ Expedited Review 45 CFR 46.110 ☐ Full Board Review 45 CFR 46

Review Date	12/12/18
IRB#	4540
Title of Project	A Flipped Classroom Approach in Undergraduate Anatomy and Physiology: A Mixed Methods Study
Researcher/s	Scott Gaines, Dr. John Williams, Mary von dem Bussche

☑ **APPROVED**

Effective duration of IRB Approval: 12/12/18 to 12/11/19

For Exempt Approved, Please Note: *while your project is exempt from providing Informed Consent information to the IRB, your project must still obtain participants' informed consent.*
For Expedited and Full Board Approved, Please Note:
a.The IRB's approval is only for the project protocol named above. Any changes aresubject to review and approval by the IRB.
b.Any adverse events must be reported to the IRB.
c.An annual report or report upon completion is required for each project. If the project isto continue beyond the twelve month period, a request for continuation of approvalshould be made in writing. Any deviations from the approved protocol should be noted.

☐ **NEEDS REVISION AND RESUBMISSION**

☐ **NOT APPROVED**

Printed Name IRB Reviewer Eugene P. Kim, Ph.D.

Signature of IRB Reviewer ____________________

Appendix K Open Codes

- Amount of information
- Anatomy
- Application
- Beneficial
- Boring
- Case Studies
- Challenging
- Check understanding
- Class Questions
- Clinical Connection
- Concept Maps
- Confidence
- Confusing
- Critical Thinking
- Detailed information
- Differing Points of View
- Dislike
- Does not prefer
- Drawing Anatomy
- Easy
- EdPuzzles
- Efficient
- Enjoyment
- Fear for Grade
- Feeling Lost in Class
- Flipped Classroom
- Frustrated
- Grades Improved
- Grading
- Group Work
- Hard to focus
- Helpful
- Improved Study Habits
- Inefficient
- Instructor Emphasis
- Instructor Interactions
- Instructor Videos
- Interest in subject
- Involvement
- Labs
- Lack of Interaction
- Lack of time
- Late viewing of videos
- LearnSmarts
- Liked
- Maturing
- Memorization
- Metacognition
- Needed
- Not applicable
- Not Engaged
- Note Packets
- Notes Fill-in-the-blanks
- Notetaking
- Optional
- Physiological Processes
- POGILs
- Practicals
- Preference
- Pressure
- Problem Solving
- Questions answered in class
- Reading Text
- Redundant
- Repetition of Material
- Retention of Material
- Scared to ask questions
- Second Guessing Themselves
- Self-paced
- Self-Sufficiency
- Slow
- Stopped Procrastinating
- Stress
- Struggle
- Student Cohesiveness
- Study alone
- Study guide
- Study in Groups
- Summary notes
- Talking in class
- Teaching each other
- Terminology
- Test Corrections
- Test Prep
- Textbook
- Time Consuming
- Traditional Lecture
- Tutoring
- Understanding Mistakes
- Video Lectures
- Virtual lab
- Visual Learner
- Visualizations 3D
- YouTube Videos

Appendix L Willowed Codes

Code	Brief Definition
Active_Learning__POGILs	Process Oriented Guided Imagery Learning
Active_Learning_Case_Studies	Case Studies
Active_Learning_Concept_Maps	Concept Maps
Active_Learning_Drawing	Drawing of anatomy in class
Active_Learning_Virtual_Labs	Virtual Labs
Amount_of_time	Perceived content volume to be considerable and time consuming
Application	Application of content to clinical settings
Boring	Student reported aspect of class boring and/or redundant
Challenging	Found specific course aspect challenging
Change(-)	Sees change as a negative
Change(+)	Sees change as a positive
Class_Grade	Class Grade converted to GPA
Confidence(-)	Student felt confidence in knowledge of material decreased
Confidence(+)	Student felt confidence in knowledge of material increased
Cum_GPA_Prior	Cumulative GPA of student prior to class
Easy	Found specific course aspect easy
Efficient(-)	Student felt aspect of the class was inefficient
Efficient(+)	Student felt aspect of the class was efficient
Engaged(-)	Student reported feeling engaged in classroom
Engaged(+)	Student reported feeling engaged in classroom
Exam	Class Exams
F_Student	Flipped Student
Fast	Content presented too fast
Flipped(-)	Did not like flip
Flipped(+)	Liked flip
Flipped (_)	Neutral or did not express opinion on flip
Frustrated	Student reported being frustrated with presentation of material
Gender	Male or Female
Grade/Content	Student emphasizes grades over understanding content
Grades(-)	Grade worsened
Grades(+)	Grade improved
Group_1	First semi-structured interview group
Group_2	Second semi-structured interview group
Incomplete_EdPuzzle	Number of incomplete/late EdPuzzles
Individualization – CUCEI_Scale	Based on Moos's Scheme: System maintenance and change dimensions
Innovation – CUCEI_Scale	Based on Moos's Scheme: System maintenance and change dimensions
Interactions_classmates	Interactions between students

Interactions_instructor	Interactions between students and instructor
Interest(-)	Student reported a decrease in interest in subject
Interest(+)	Student reported an increase in interest in subject
Intimidated_class_environment	Student felt intimidated by classroom environment
Intimidated_material	Student felt intimidated by material
Involvement–CUCEI_Scale	Based on Moos's Scheme: Relationship Dimensions
Labs	Laboratory exercises
Lecture_Discussion(-)	Discussion during lecture increased
Lost	Student felt lost/confused in class
Memorization	Memorization stressed as a study habit/skill
Notetaking	Student expressed notetaking strategies
Notetaking_note_packets	Instructor provided note packets coinciding with lecture
Ownership	Ownership of the learning process
Personalization – CUCEI_Scale	Based on Moos's Scheme: Relationship Dimensions
Practicals	Laboratory Practicals
Problem_Solving	Problem solving and critical thinking necessary in aspects in class
Procrastination(-)	Procrastination increased
Procrastination(+)	Procrastination decreased
Reading_text	Student read assigned pages from text
Repetition(+)	Hearing topics a second time helped learning
Repetition	Repetition of material
Satisfaction–CUCEI_Scale	Based on Moos's Scheme: Relationship Dimensions
Scientific_misconceptions	Scientific misconceptions discussed in class
Slow	Content presented too slowly
Student_Cohesiveness – CUCEI_Scale	Based on Moos's Scheme: Relationship Dimensions
Study_groups(-)	Student reported preferring studying alone
Study_groups(+)	Student reported benefits/preferences of studying in groups
Study_Habits_good	Explicitly states good study habits
Study_Habits_poor	Explicitly admits to poor study habits
Study_Habits(-)	Feels Study habits worsened
Study_Habits(+)	Feels Study habits improved
T_Student	Traditional Student
Traditional(-)	Did not like Traditional Class
Traditional(+)	Liked Traditional Class
Video_lectures_late	Student reported viewing video lectures after deadline
Video_lectures(-)	Disliked video lectures

Video_lectures(+)	Liked video lectures
Visual_Learner	Student considers themselves to be a “visual learner”

Appendix M Sample Note Packet

1 The Endocrine System

Overview of the Endocrine System

- Endocrine vs Exocrine Glands
- Endocrine System vs Nervous System
- Hormones

Endocrine vs Exocrine Glands

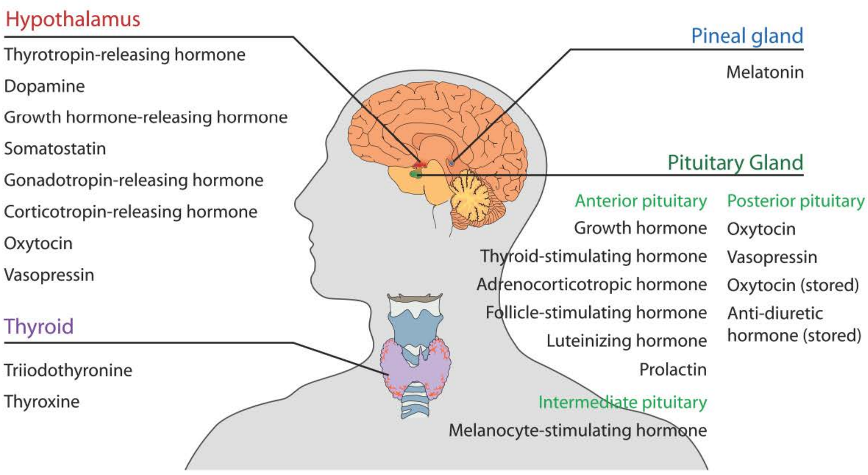
Endocrine System vs Nervous System and Other forms of Intracellular Communication

Classes of Hormones

Mechanisms of Hormone Action

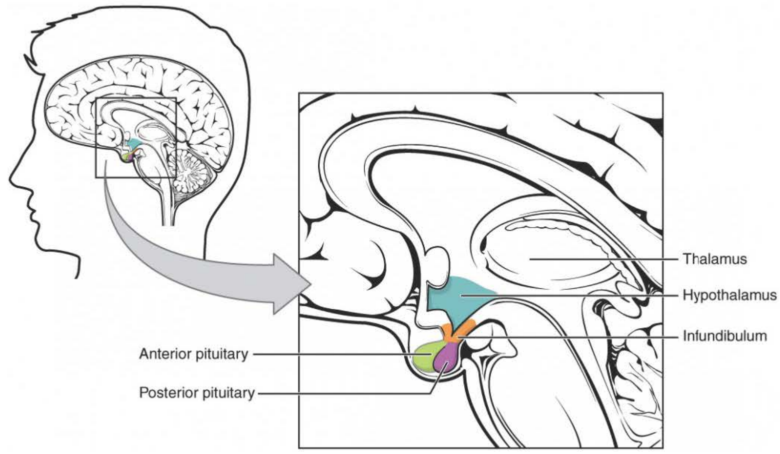
What Stimulates Hormone Secretion (i.e. "Endocrine Reflexes")

Hormone Clearance??



The Hypothalamus and Pituitary Gland (H-P Axis)

Anatomy of the Hypothalamus and Pituitary Gland



Embryologic Development of Pituitary

The Hypothalamus

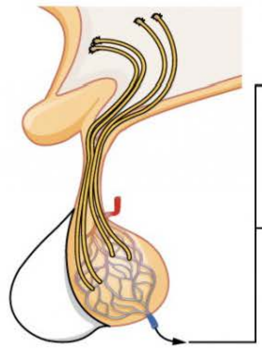
- Acts as an ____________________ organ (ADH and oxytocin – stored in post pituitary)
- Secretes ____________________ ____________________ that control endocrine cells in pituitary
- Contains autonomic ____________________ centers that exert direct control over endocrine cells of adrenal medulla

The Pituitary

- Also called ____________________
- Bilobed gland (with anterior and posterior lobes)
- Lies within sella turcica
- Hangs inferior to hypothalamus
  - Connected by ____________________
- Releases ___ important peptide hormones (well, at least 8 important peptide hormones)
  - 7 from the Anterior Pituitary
  - 2 from the Posterior Pituitary (from the hypothalamus)
- Hence, use Second Messenger System (Hormones bind to membrane receptors)
  - Use cAMP as second messenger
- Of the 9 important peptide hormones, 4 are ____________________ hormones
  - So, what's a tropic hormone?
  - Tropic Hormones "turn on" endocrine glands to secrete other hormones

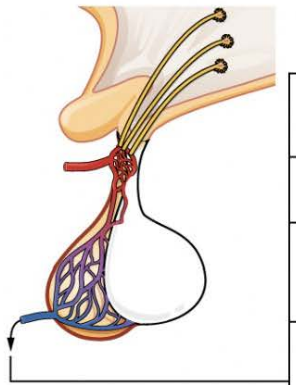
Anterior Pituitary	Posterior Pituitary
Tropic Hormones	**1.** ADH (Antidiuretic Hormone) ♦ Made by supraoptic nuclei of hypothalamus
1. TSH (Thyroid-stimulating Hormone or thyrotropin)	
2. ACTH (Adrenocorticotropic Hormone)	
The Gonado<u>tropins</u>	**2.** OXT (Oxytocin) Made by paraventricular nuclei of hypothalamus
3. FSH (Follicle-stimulating Hormone)	
4. LH (Luteinizing Hormone)	
Direct Hormones	
5. Growth Hormone	
6. Prolactin	
7. Melanocyte-Stimulating Hormone	

Posterior Pituitary Hormones

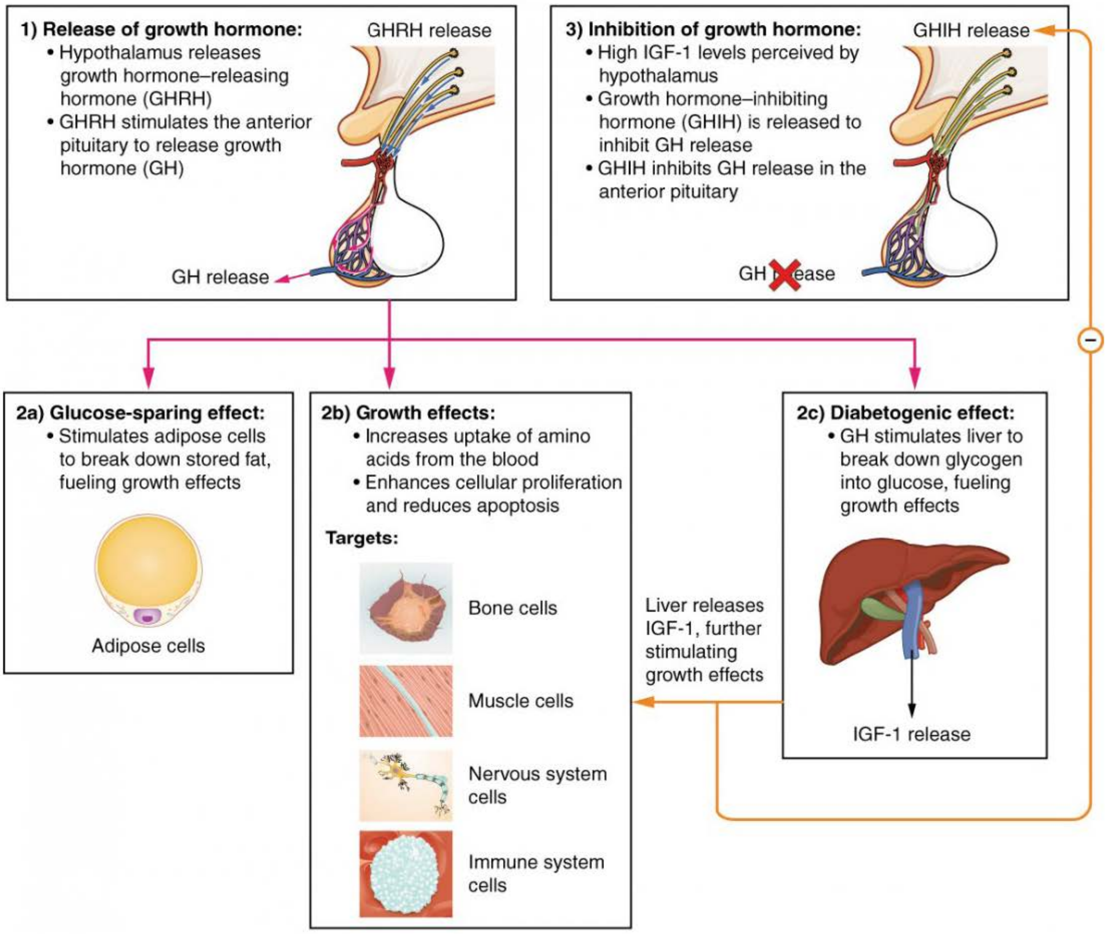


Releasing hormone (hypothalamus)	Pituitary hormone	Target	Effects
ADH	Stores ADH	Kidneys, sweat glands, circulatory system	Water balance
–	OT	Female reproductive system	Triggers uterine contractions during childbirth

Anterior Pituitary Hormones



Releasing hormone (hypothalamus)	Pituitary hormone	Target	Effects
GnRH	LH	Reproductive system	Stimulates production of sex hormones by gonads
GnRH	FSH	Reproductive system	Stimulates production of sperm and eggs
TRH	TSH	Thyroid gland	Stimulates the release of thyroid hormone (TH). TH regulates metabolism.
PRH (inhibited by PIH)	PRL	Mammary glands	Promotes milk production
GHRH (inhibited by GHIH)	GH	Liver, bone, muscles	Induces targets to produce insulin-like growth factors (IGF). IGFs stimulate body growth and a higher metabolic rate.
CRH	ACTH	Adrenal glands	Induces targets to produce glucocorticoids, which regulate metabolism and the stress response

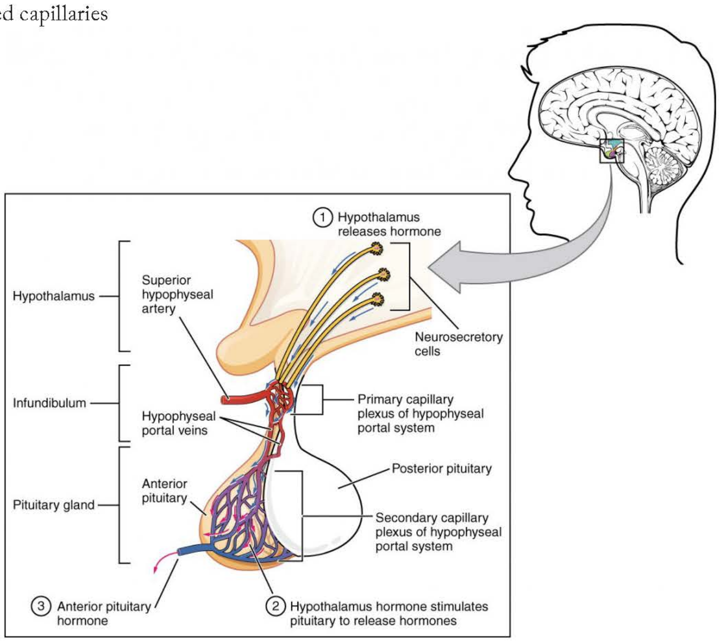

1) Release of growth hormone:
• Hypothalamus releases growth hormone–releasing hormone (GHRH)
• GHRH stimulates the anterior pituitary to release growth hormone (GH)
GHRH release
GH release
3) Inhibition of growth hormone:
• High IGF-1 levels perceived by hypothalamus
• Growth hormone–inhibiting hormone (GHIH) is released to inhibit GH release
• GHIH inhibits GH release in the anterior pituitary
GHIH release
2a) Glucose-sparing effect:
• Stimulates adipose cells to break down stored fat, fueling growth effects
Adipose cells
2b) Growth effects:
• Increases uptake of amino acids from the blood
• Enhances cellular proliferation and reduces apoptosis
Targets:
Bone cells
Muscle cells
Nervous system cells
Immune system cells
Liver releases IGF-1, further stimulating growth effects
2c) Diabetogenic effect:
• GH stimulates liver to break down glycogen into glucose, fueling growth effects
IGF-1 release

The Anterior Lobe of the Pituitary Gland

- Also called ___________________
  - Hormones "turn on" endocrine glands or support other organs
  - Has three regions
    - Pars distalis
    - Pars tuberalis
    - Pars intermedia

The Hypophyseal _______________ System

- Median eminence
  - Swelling near attachment of infundibulum
  - Where hypothalamic neurons release regulatory factors
    - Into interstitial fluids
    - Through fenestrated capillaries
- ___________________ Vessels
  - Blood vessels link two capillary networks
  - Entire complex is portal system
    - Ensures that regulatory factors reach intended target cells before entering general circulation



Hypothalamic Control of the Anterior Lobe

- Two classes of hypothalamic regulatory hormones
    - ____________________ hormones (RH)
        - Stimulate synthesis and secretion of one or more hormones at anterior lobe
    - ____________________ hormones (IH)
        - Prevent synthesis and secretion of hormones from the anterior lobe
    - Rate of secretion is controlled by negative feedback

The Posterior Lobe of the Pituitary Gland

- Also called ____________________
    - Contains unmyelinated axons of hypothalamic neurons
    - Supraoptic and paraventricular nuclei manufacture:
        - ____________________ hormone (ADH)
        - ____________________ (OXT)