Assessing High School Biology Academic Achievement by Comparing Traditional Versus Virtual Dissection of Rat Specimens

Arthur L. Jamison Jr
Lynn University

Follow this and additional works at: https://spiral.lynn.edu/etds
Part of the Science and Mathematics Education Commons

Recommended Citation
https://spiral.lynn.edu/etds/145

This Dissertation is brought to you for free and open access by the Student Work at SPIRAL. It has been accepted for inclusion in Student Theses, Dissertations, Portfolios and Projects by an authorized administrator of SPIRAL. For more information, please contact liadarola@lynn.edu.
ASSESSING HIGH SCHOOL BIOLOGY ACADEMIC ACHIEVEMENT BY COMPARING TRADITIONAL VERSUS VIRTUAL DISSECTION OF RAT SPECIMENS

By

Arthur L. Jamison, Jr.

A Dissertation in Practice

Submitted in partial fulfillment of the requirements for the degree of

Doctorate in Education

Doctoral Program in Educational Leadership at Lynn University,
College of Education

Lynn University
2014
ASSESSING HIGH SCHOOL BIOLOGY ACADEMIC ACHIEVEMENT BY COMPARING TRADITIONAL VERSUS VIRTUAL DISSECTION OF RAT SPECIMENS

Abstract

American's lagging student performance on standardized assessments in critical subject areas such as science has in-part led to numerous educational reforms including No Child Left Behind and Race to the Top, resulting in an increased focus on student achievement. International and national government officials note the direct correlation between the value of science and driving the economic prosperity of a nation adding increased pressure to improve science scores in the United States. Local districts and schools struggle with how to improve student achievement in order to meet the requirements of state and federal educational reforms.

The purpose of this mixed-methods study was to examine the learning effectiveness through the work of implementing rat dissection for comparison between the traditional hands-on method and the alternative virtual simulation. Student achievement was measured by pre-test, posttest assessments and a lab practicum, which measured students' recall knowledge on rat anatomy. Students' attitudes were measured by a post-dissection attitude survey.

The sample consisted of 311 high school biology students from 20 classes taught by three instructors. The school had a student enrollment of approximately 2400 with a minority population of 70%. Prior to the activity, each teacher administered a pre-test assessment to the classes. Next, each teacher divided the classes in half. One group conducted the hands-on dissection activity on a rat, and the other group completed the Rat Dissection 1.1 virtual program activity. Each group then took a posttest assessment and completed the lab practicum. Subsequently, the groups switched and completed the opposite activity. Finally, a survey was administered.

The data was examined using an independent samples t test and a MANOVA model. Results indicated that students who participated in the hands-on dissection activity made significantly larger gains on the post-test assessment, but not on the lab practicum. Student attitude was also analyzed by calculating frequencies of the survey questions. Results indicated a preference of two to one in favor of the hands-on dissection method.
ASSESSING HIGH SCHOOL BIOLOGY ACADEMIC ACHIEVEMENT BY
COMPARING TRADITIONAL VERSUS VIRTUAL DISSECTION OF RAT
SPECIMENS

Jamison, Jr., Arthur L., Ed. D.

Lynn University, 2014

© 2014 by Jamison, Jr., Arthur L.
ALL RIGHTS RESERVED

U.M.I.
789 Eisenhower Parkway
Ann Arbor, MI 48106-1346
ACKNOWLEDGEMENTS

First I would like to thank my Lord and Savior, Jesus Christ, for giving me the strength and perseverance to complete this journey.

Much love and appreciation to my wife, Mishka Lisa-Marie Symonette-Jamison

Gratitude to my family for their support, and inspiration: Arthur Sr., Paula Cruz (Mom) and Aixa (Sister)

Special thanks to the St. Lucie County School District team: Mrs. Kerry Padrick, Dr. Kathleen Daily and Mr. Todd Smith (Principal Fort Pierce Central)

Many thanks to Dr. William Leary, Dr. Adam Kosnitzky, Dr. Kevin G. Perry and Dr. Craig Mertler for guidance, counseling and professional mentorship

For those who continued to say I could do it, pushing me out of my academic and scholarly comfort zone with thoughtful and artful dialogue: Susan Seal Shari Ashe Maurice Bonner Dr. Pressley Charles Dr. Rodney Harley (Brother) James Williams (Brother) Jennifer Kadament Holly O’ Brien Crystal Woodard

Most notably: Dr. Kerryanne Monihan – For recommending the Lynn University Doctoral Program
# TABLE OF CONTENTS

Chapter 1: Introduction ................................................................................................................. 1
  Purpose of Study ......................................................................................................................... 1
  Statement of the Problem .......................................................................................................... 1
  Research Questions ..................................................................................................................... 2
  Background ................................................................................................................................ 3
  Definition of Terms ..................................................................................................................... 7
  Rationale of the Study ................................................................................................................. 9
  Assumptions ................................................................................................................................ 9
  Scope, Limitations, and Delimitations ....................................................................................... 9

Chapter 2: Literature Review ...................................................................................................... 11
  Student Achievement in Science ............................................................................................... 13
  Biology ....................................................................................................................................... 17
  History of Dissection .................................................................................................................. 18
  Organism Usage Statistics for Dissection .................................................................................. 20
  Purpose of Dissection .................................................................................................................. 22
  Proponents for Use of Hands-On Dissection .......................................................................... 25
  Opponents against Hands-On Dissection ................................................................................. 31
  Health and Environment Arguments ...................................................................................... 37
  Chemicals ................................................................................................................................. 38
  Diseases ..................................................................................................................................... 39
Mortality ................................................................. 57
Hawthorne Effect .................................................. 57
Threats to External Validity ...................................... 57
Interactive Effects of Testing .................................... 57
Interactive Effects of Selection Biases ....................... 58
Multiple Treatment Interference .............................. 58
Ethical Considerations ........................................... 58
Summary ................................................................. 59

Chapter 4: Data Analysis ........................................... 60
Descriptive Statistics .............................................. 60
Research Question 1 ................................................ 69
  Independent samples $t$ test .................................... 70
  MANOVA model .................................................... 72
Research Question 2 ................................................ 75
  Independent samples $t$ tests, differences by gender .... 75
  MANOVA analysis, differences by gender ................ 77
  Independent samples $t$ test, differences by ethnicity ... 79
  MANOVA analysis, differences by ethnicity ............... 86
Research Question 3 ................................................ 88
Summary ................................................................. 90
Chapter 5: Discussion ................................................................. 92
Summary of Findings ................................................................. 92
Conclusions .................................................................................. 96
Limitations ................................................................................... 98
Recommendations for Future Research ....................................... 100
Implications for Practice .............................................................. 103
References .................................................................................... 108
Appendix ....................................................................................... 124
LIST OF FIGURES

Figure 1. Comparison of Score of Florida 4th Graders to the Nation on NAEP. ............ 14

Figure 2. Comparison Score of Florida 8th Graders to the Nation on NAEP. ............ 15

Figure 3. Percentage of Change in State Funding since 2008 (Center on Budget Policies and Priorities, 2013). .............................................................................................................. 40

Figure 4. Type of instruction received by students in the sample. ............................ 60

Figure 5. Gender breakdown of sample. Note the sample is close to even, with slightly more male students than female students................................................................. 60

Figure 6. Breakdown of sample by ethnicity. Note that the sample is fairly equally divided between Caucasian, African-American, and Hispanic students. .......................... 61

Figure 7. Mean scores of all students for the pretest, posttest, and practicum. .......... 62

Figure 8. Pretest, posttest, and practicum scores for all students under both the traditional and virtual dissection methods................................................................. 63

Figure 9. Pretest, posttest, and practicum scores for all students broken down by gender. .................................................................................................................................. 64

Figure 10. Pretest, posttest, and practicum scores for all students broken down by ethnicity.................................................................................................................................. 65

Figure 11. Pretest, posttest, and practicum scores for male students only, broken down by type of instruction. ........................................................................................................ 66

Figure 12. Pretest, posttest, and practicum scores for female students only, broken down by type of instruction. ........................................................................................................ 67
Figure 13. Pretest, posttest, and practicum scores for Caucasian students only, broken down by type of instruction. ................................................................. 68
Figure 14. Pretest, posttest, and practicum scores for African-American students only, broken down by type of instruction. ......................................................... 68
Figure 15. Pretest, posttest, and practicum scores for Hispanic students only broken, down by type of instruction. ................................................................. 69
Figure 16. Mean score improvement for all students and by type of instruction. ............ 70
Figure 17. Boxplot of score improvement variable for both traditional and virtual dissection groups, where t test confirmed difference is significant. ....................... 71
Figure 18. Q-Q plot of posttest scores. Note that the distribution appears approximately normal except at both ends of the distribution. ................................. 73
Figure 19. Q-Q plot of practicum scores. Note that the distribution appears approximately normal, but not continuous because of the nature of the practicum scores. ................................................................. 73
Figure 20. Mean score improvement for virtual dissection students only, broken down by gender. Note that the mean score improvement appears significantly higher for male students. ................................................................. 76
Figure 21. Boxplot of score improvement variable for both male and female students in the virtual dissection group only. ................................................................. 77
Figure 22. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case Caucasian and non-Caucasian students. ............... 80
Figure 23. Boxplot of score improvement variable for Caucasian and non-Caucasian students in the virtual dissection group only. ................................................................. 81

Figure 24. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case African-American and non-African-American students. .............. 82

Figure 25. Boxplot of score improvement variable for both African-American and non-African-American students in the virtual dissection group only. ................................. 83

Figure 26. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case Hispanic and non-Hispanic students. ........................................... 84

Figure 27. Boxplot of score improvement variable for both Hispanic and non-Hispanic students in the virtual dissection group only. ................................................................. 85

Figure 28. Pie chart showing which type of dissection instruction students preferred. .... 88

Figure 29. Type of dissection preferred by type of instruction originally received. ...... 89
LIST OF TABLES

Table 1. Published Studies Noting Traditional Hands-on Dissection to be more Effective than Virtual Dissection. ............................................................................................................................................... 31

Table 2. Published Studies Noting Virtual Dissection to be more Effective than Traditional Hands-on Dissection. ............................................................................................................................................... 36
CHAPTER 1: INTRODUCTION

Purpose of Study

The purpose of this mixed-methods study was to examine the learning effectiveness through the work of implementing rat dissection for comparison between the traditional hands-on method and the alternative virtual simulation. With school districts seeking to increase the achievement rates in science, this study sought to determine whether presentational and activity methods would need to be addressed in increasing student results. Furthermore, this research project was designed to a) evaluate the effectiveness on academic achievement of a virtual rat dissection in the study of human systems focusing on the various organs and the functions of the organs as compared to a traditional hands-on rat dissection, and b) explore high school biology students’ affective response to dissection and its relationship towards animal use in the study of the sciences.

Statement of the Problem

As American students continue to underachieve on international standardized assessments such as the Program for International Student Exam (PISA) which assesses a student’s knowledge in the subject areas of reading, math, and science and the Trends in International Mathematics and Science Study exam (TIMMS) which assesses the subject areas of math, and science, the United States federal government continues to reform education in part emphasizing its focus on standardized test scores in evaluating the efficacy of our public school system. This has led schools to focus on how the content and activities are presented that would benefit the students’ engagement, thus leading to improved student achievement on these various standardized assessments.
Furthermore, with both sides having valid points, the controversy relating to the effectiveness of traditional dissection versus virtual dissection exists. The school under consideration in the study uses strictly traditional dissection methods. It is the intent of this study to provide an understanding of how virtual dissection software can affect a biology student's depth of knowledge thus leading to an increase in student achievement within the topic of human systems. This determination will be demonstrated through student achievement results regarding human systems on general animal anatomy assessments and a lab practicum exercise, thus aligning with the district's initiative to meet the demands of No Child Left Behind (NCLB) and Race to the Top (RttT) educational reforms.

**Research Questions**

Three research questions guided this investigation:

1. What effect on student achievement do virtual animal dissection and traditional dissection methods have on a general animal (rat) anatomy assessment (consisting of recall questions), and a lab practicum exercise (consisting of organ identification and function)?

2. What effect does virtual dissection (as opposed to the traditional method) have on student achievement focusing on gender and the primary ethnicities (Caucasian, African-American and Hispanic)?

3. What preference do high school biology students have as a means for acquiring the knowledge of human systems, virtual dissection of organisms or the traditional dissection method?
Background

According to local, state, national and international assessments in science American students underachieve in the subject of science. Against our international counterparts, data indicates that American students lag behind in reading, math, and science (Fleishman Hopstock, Pelczar, & Shelley, 2010; Gonzales, Brenwald, Jocelyn, Kastberg, Roey & Williams, 2008). Educational reforms such as No Child Left Behind (U.S. Department of Education, 2004) and the current Race to the Top (The White House, 2009) legislation have, in part, been driven by the dismal student performance on these assessments. Nationally, science innovation and technology play a vital role in the economic stability of the United States and its competitive countries, driving the need for educational institutions to provide a skilled workforce. Furthermore, the Organization for Economic Co-Operation and Development (OECD) stated that, “science, technology, and innovation are now keys to improving a country’s economic performance and social well-being” (OECD Observer, 2000).

At the local level, there is also a sense of urgency for districts to raise student achievement results to meet the demands of No Child Left Behind (NCLB) and Race to the Top (RttT) (Kober, Chudowsky, & Chodowsky, 2010). To meet these demands, many school districts throughout the state of Florida implemented benchmark tests as an instrument to evaluate student progress and drive instruction. The urgency to increase the achievement rates in science has led to the question, how are the concepts being presented in the specific subjects within the science curriculum? In answering the question, biology emerges to the forefront. It is the subject that is assessed by the state of Florida at the high school level for science (End of Course exam [EOC]). Also, it is the
typical age (15 years old) that American students are tested and measured versus our international counterparts on many science content exams.

Within the course of high school biology there are six standards or topics that encompass the curriculum: Nature of Science, Organization and Development of Living Organisms, Division and Evolution of Living Organisms, Heredity and Reproduction, Interdependence and Matter and Energy Transformations (Bybee, 2002). In high school biology, the study of human systems within the national biology standard of Organization and Development of Living Organisms is covered during the third quarter of the school year. While covering the topic of human systems, it is customary for a teacher to cover the various organs and their functions within the scope of the material. In many cases, instructors, use animal cadavers through the process of dissection, as a model representation of the various human body systems and their functions.

The use of animals for dissection has been a foundational practice in American colleges since the 1800’s and in secondary schools since the 1920’s (Barr & Herzog, 2001; Ethical Science and Education Coalition [ESEC], 2001). Great controversy has also followed this teaching practice since its inception. Moral, ethical, and religious objection to this practice and the realization that teacher dissection practices have the potential to greatly impact the student learning experience have resulted in increased attention to alternative methods of teaching biology (Almy, Goldsmith, & Patronek, n.d.; Balcombe, 1997; Cunningham, 2000).

The 1980’s was a time of intense debate with regard to the advantages and disadvantages of dissection as a teaching practice (Balcombe, 2000; Gilmore, 1991; Keiser & Hamm, 1991; Offner, 1993; Orlans, 1988; Strauss & Kinzie, 1991). This
debate has raised awareness about student concerns, influenced the development of innovative instructional tools, and extended academic discussions toward the philosophical realm. Because of these reasons, the introduction of virtual technological platforms has become increasingly marketed as an alternative for traditional animal dissections.

Proponents of traditional dissections believe that no simulation can replace the experience of the sight, smell, and feel of actual animal tissue (Morrison, 1992; Russo, 1997; Valli, 2001). Furthermore, these educators feel that computer simulations are identical for every user; however, life and traditional dissection are not. Traditional dissection, therefore, offers students the chance to see life as it is, in its pure form, whereas alternatives do not (Russo, 1997).

Others in favor of traditional dissection believe that there are other lessons to be learned aside from teaching anatomy. Organizations such as the National Science Teachers Association (NSTA) (2005), the American Medical Association (AMA) (Riechard, 1993), and the United States Military (Anderson, 1992) all support traditional dissection. They believe that it is an essential element of a student’s education. It affords students the benefit of the process skills associated with science (Berstein, 2000) and Darwin’s concept of variation among species (Russo, 1997), further contributing to many medical breakthroughs in the past.

Opponents of dissection note many reasons for their view as well. One such reason is that in order for true learning to take place, repetition is necessary. The inherent problem with dissection is its destructive rather than constructive process, which destroys many of the specimen’s structures and their spatial relations, precluding reexamination by
the student (Rose, 1995). With virtual (computer) simulations, a dissection can be repeated a number of times to ensure learning (Hepner, 1996). Additionally, because traditional dissection is widely believed to be the status quo, opponents' state that teachers instruct their students in the manner that they were taught; it is what they are comfortable with, so the practice passes on. Conversely, with the increased realism of computer imaging, opponents stand firm in their belief that the practice of dissection should be eliminated from the biology curriculum in its entirety.

Moreover, every year millions of animals are dissected in elementary, secondary, and college science classes (People for the Ethical Treatment of Animals [PETA], n.d.). The Humane Society of the United States (HSUS) (2004) claimed that most of the dissection organisms were not raised in laboratories but rather were collected from the wild. The Humane Society also claimed that this practice has contributed to the decline of wild animal populations.

Other concerns that opponents present regarding traditional dissections are the spread of disease and the chemicals that are used by the industry. Although there have been no documented cases of diseased animals used for dissection in the United States, China has found the rabies virus in animals imported for laboratory use (Sharma, 2006). This is concerning due to the fact that high school biology teachers usually purchase laboratory specimens from supply vendors. The process of preparing the specimen(s) to be used in the science market involves harsh chemicals used as preservatives to stabilize the cadaver for long term storage. These chemicals may cause skin and or respiratory irritations to those who are exposed to them (Jackson, 1991).
Definition of Terms

The terms pertinent to this study include:

**Achievement rates.** These are the degree of performance in different subject areas in various levels of education (Department of Education, 2006).

**Benchmarks.** The battery of tests administered by the participating school district, used for progress monitoring (formative Benchmarks) and of mastery summative exams (St. Lucie County School Board [SLCSB], 2012).

**Content knowledge.** This refers to knowledge specific to the course or topic and defined by the standards for that grade level course (Gonzales, 2001).

**End of course assessments.** The end of course examinations that are computer-based, criterion-referenced assessments that measure the Next Generation Sunshine State Standards for specific courses, as outlined in their course descriptions (Florida Department of Education [FLDOE], n.d.).

**Human systems.** This refers to the concept that organ systems are composed of two or more different organs that work together to provide a common function (Carpi, 1999).

**Traditional dissection.** This is the hands-on method of observing or cutting into a dead animal for the purposes of learning anatomy or physiology (Ethical Science & Education Coalition, 2001); to cut apart for scientific examination, usually in reference to the study of animals or humans (John Hopkins Bloomberg School of Public Health, 2009).
Virtual dissection. Alternatively, this is the interactive dissection of an animal by means of a computer used to imitate the process as performed on a real organism (Akpan & Andre, 2000).
Rationale of the Study

There have been a few studies that examine the infusion of virtual dissections in high school biology classrooms. Many of these studies have focused on virtual frog dissections (Apkan & Andre, 2000; Kinzie, Strauss, & Foss, 1993; Sweitzer, 1996). This study used rat dissection specimens, as rats are believed to provide the students a closer analysis of an anatomy to that of the human system. Furthermore, the results of this study could provide an immediate solution to some practical problems that exist in today’s high school biology classrooms. Additionally, with respect to the research community, the researcher has found no investigative evidence denoting that a virtual rat dissection program has been used within a high school biology setting. Therefore, this study should provide high school biology instructors new information on student learning and achievement and its relation to the two presentational approaches.

Assumptions

The following assumptions were made by the researcher at the onset of this study:

1. Students passed their previous science course (physical science) in the ninth grade.
2. The teacher-created assessments and lab practicum accurately assess student content knowledge in biology.
3. Participating teachers possess the knowledge and skills necessary to appropriately deliver the content and implement the activity.

Scope, Limitations, and Delimitations

The sample was limited to a single, mid-size high school in a district located in Central-South Florida. The biology teachers at this high school participated voluntarily.
Therefore, there could be a school and/or teacher effect. An effect to this study could be the variance in time that the study is conducted from teacher to teacher based on the teacher’s place in covering the curriculum leading to the sharing of information. This means that the participants (students) were not in a controlled environment thus the study cannot reveal a correlation between time and the test results. Also, since animal dissection presentations can be implemented in a variety of ways, the teachers were encouraged to deliver the information deemed most appropriate for their students. Consequently, this meant that there was not one approach to the presentation of the material and the study cannot reveal connections between student achievement and mechanisms of implementation.

Secondly, the high school students were enrolled in either a Biology I or Honors Biology I course. The student subjects were diverse in both gender and race. Results should be generalizable to similar populations, but not necessarily to students, schools, or districts that vary substantially from the sample population of this study.

Another limitation of this study is that the students were enrolled in a suburban public high school. The results from this study may not be used to generalize to those obtained from a public inner city high school or private high school setting. Furthermore, the researcher found no research evidence denoting that a virtual rat dissection program has been used within a high school biology setting.
CHAPTER 2: LITERATURE REVIEW

“But in many ways, our future depends on what happens in those contests, what happens when a young person is engaged in conducting an experiment, or writing a piece of software, or solving a hard math problem, or designing a new gadget. It's in these pursuits that talents are discovered and passions are lit, and the future scientists, engineers, inventors, entrepreneurs are born. That's what's going to help ensure that we succeed in the next century, that we're leading the world in developing the technologies, businesses and industries of the future.

And this is the reason my administration has put such a focus on math and science education, because despite the importance of inspiring and educating our children in these fields, in recent years the fact is we've been outpaced by a lot of our competitors. One assessment shows that American 15-year-olds ranked 21st in science and 25th in math when compared to their peers around the world.

......It is unacceptable to me, and I know it's unacceptable to you, for us to be ranked on average as 21st or 25th, not with so much at stake. We don't play for second place here in America. We certainly don't play for 25th.” (Obama, White House Science Fair Speech, 2010)

The Organization for Economic Co-operation and Development (OECD) indicated that the reduction on the reliance of natural resources and a focus on science, technology and informational knowledge are the keys to improving a country’s economic performance and social well-being (OECD, 2011). Countries such as Finland, Japan, South Korea, New Zealand, and Singapore who lack natural resources that can drive a country’s economy understood that as a nation, they can prosper by the way of
technological advancements and what comes out of the minds of their citizens. Leaders from these countries undertook important curricular reforms, and educators were given more freedom to experiment (The New York Times, 2010) which resulted in some of the highest scores per country on international assessments.

These results were alarming to the United States. Educational leaders and government officials saw this as an issue because this is what OECD identified as our country’s economic boom in the late 1990’s; driven by scientific and technological advances through education (OECD, 2000) which helped stimulate the United States’ economy. With recent history in mind, decision makers forecasted remarkable economic growth for specific competitor countries to that of the United States thus affecting our country’s economic upswing.

Furthermore, the United States had witnessed an economic period such as this once before. It was the reason that the National Commission on Excellence in Education published *A Nation At Risk*. Due in part to the country of Russia’s economic growth spearheaded by their technological advances in the 1960’s thru 1970’s, the report indicated that American students were being out-competed by our international counterparts and without instantaneous and drastic improvements; the country was headed towards economic peril.

In 2008 United States leaders convened at Princeton University determined to seek a solution(s) relating to the impending crisis. The data used was the comparison of competing industrial nations to the United States on science achievement assessments such as the Programme for International Assessment (PISA) and the Trends in Mathematical and Science Study (TIMSS). Both tests assessed the knowledge of 15-
year-olds in science and mathematics. According to the results, it showed that the students in America had not demonstrated progress calling for a renewed commitment to science and technology.

Subsequently, the purpose of this study was to investigate the impact of a student activity designed to improve science content knowledge. First, I will present more detail information on student achievement in science. Second, I will associate the subject of biology and the importance that it plays in science student achievement in the United States. Third, I will discuss biology content matter and the role that dissection activities have in teaching the subject matter. Fourth, I will elaborate on the history of dissection and present views from both sides – those for and against dissection of animals for science research. Finally, I will elaborate two points, the cost difference between both methods and safety issues that arise when performing a dissection activity.

**Student Achievement in Science**

Nationally, the achievement results of U.S. students in science has been stationary and/or in a declining state on multiple assessments including the National Assessment of Educational Progress (NAEP), the Program for International Student Assessment (PISA) and Trends in Mathematics and Science Study (TIMSS). Compounding the issue, there is a significant gap in science achievement with African-American and Hispanic students considerably underperforming against White and Asian/Pacific Islander students (National Center for Educational Statistics, 2011). Students in inner-city school districts are outperformed against their counterparts in suburban schools, and on average boys performed slightly better than girls across categories in science.
Across the nation, the NAEP was administered in 2009. It measured students' knowledge of physical science, life science, and the Earth and Space sciences. The results revealed, just 21% of 12th graders, 30% of 8th graders and 34% of 4th graders scored at or above proficient with less than 2% scoring at the advanced level in any grade. Particularly relevant to this study, Florida students in 4th grade scored on average with the rest of the nation, but in 8th grade science scores were below the national average (National Center for Educational Statistics, 2011). See Figure 1 and Figure 2.

Figure 1. Comparison of Score of Florida 4th Graders to the Nation on NAEP.

Historically, the NAEP science assessment occurs every four years, the last science administration was in 2009. The 2011 NAEP science assessment at grade 8 was a special administration to permit comparisons with the Trends in International Mathematics and Science Study (TIMSS). The next assessment of science at grades 4, 8, and 12 is slated for 2015 to align with TIMSS henceforth.
The most recent documentation of the PISA results in 2009 was designed to assess the scientific literacy of students. American students typically at the age of 15 years old are tested, and measured versus our international counterparts on many science content exams. Focusing on age 15 provides an opportunity to measure broad learning outcomes, while all students across the many participating nations are still required to be in school (National Center for Educational Statistics, 2010). The PISA assessment not only measures content knowledge, but also a student’s ability to explain phenomena, to draw evidenced-based conclusions and the awareness of how science and technology impact and shape our society (Lau, 2009). The results showed an average score for all countries of 500 in 2009 and U.S. students scored an average of 502 placing it behind many non-Organization for Economic Cooperation and Development (OECD) countries such as Singapore, Liechtenstein, Hong Kong, and Macao, with a ranking of 17th amongst all OECD countries. Positively, the 2009 score was an improvement for U.S.
students in comparison to 2006 when the average score was below the OECD average (Highlights from Pisa, 2009).

Another international assessment, the Trends in Mathematics and Science Study was administered in 2011. It is designed to assess both content knowledge, and cognitive skill level (knowing, applying, and reasoning). In this assessment, U.S. students in 4th and 8th grade scored above the average for participating countries with scores of 544 and 525 compared to an average score of 500 (International Association for the Evaluation of Educational Achievement, 2011). It is to be noted that several countries that the U.S considers to be competitors economically outperformed the United States on the TIMSS including Singapore, Korea, Japan, Russian Federation, England, and Israel.

The Florida Comprehensive Assessment Test (FCAT) in Science was administered in 8th grade and 11th grade. The trend in performance at both grade levels since initial implementation has risen in scores. Focusing on the high school level, student scores in 11th grade have incrementally risen each year starting at 33% scoring proficient or better in 2003 to 40% in 2011 (Florida Department of Education, 2013).

In 2012, the Florida Department of Education (FLDOE) shifted its focus from general subject matter assessments (FCAT) to subject specific end-of-course exams (EOC). Included in the first wave of subject matter assessments is Biology. After the first administration of the Biology EOC student achievement remained low across the board with a state average passing rate of 50%. An even more ominous trend became evident when scores are disaggregated. While 50% of White students scored proficient or better in 2012, just 45% and 42% of Hispanics and African-Americans did so respectively in 10th grade (FLDOE, 2013). Additionally, there is wide disparity among
districts in terms of biology student achievement. Districts such as Gilchrist and Wakulla counties averaged a passing score of 53% for 10th graders that took the assessment. While Glades County School District had a 39% average passing rate. The district in this study fell slightly below state average with 10th graders at 48% proficiency or better (FLDOE, 2013).

Biology

The urgency to increase the achievement rates in science has led to the question, how are the concepts being presented in the specific subjects within the science curriculum? In answering the question, biology emerges to the forefront. As mentioned before, it is the subject that is assessed by the state of Florida at the high school level for science (End of Course exam [EOC]). The Florida EOC Assessments are part of Florida's Department of Education's strategic plan for the purpose of increasing student achievement and improving college and career readiness. EOC assessments are computer-based, criterion-referenced assessments that measure the Next Generation Sunshine State Standards for specific courses, as outlined in their course descriptions (FLDOE, 2013). Beginning in the 2013-2014 school year, current and future students in a biology class will have to pass the EOC as part of their graduation requirements.

Within the course of high school biology there are six standards or topics that encompass the curriculum: Nature of Science, Organization and Development of Living Organisms, Division and Evolution of Living Organisms, Heredity and Reproduction, Interdependence and Matter and Energy Transformations (Bybee, 2002). These six topics are the bodies of knowledge within the national biology standards that a high school biology teacher should cover over the course of a year.
In high school biology, the study of human systems within the national biology standard of Organization and Development of Living Organisms is covered during the third quarter of the school year. While covering the topic of human systems, a teacher would cover the various organs and their functions within the scope of the material. In many cases, instructors, use animal cadavers through the process of dissection, as a model representation of the various human body systems and their functions.

History of Dissection

Although the field of science has progressed rapidly, little change in content have come about in high school biology (Mayer, 1982). Textbooks of today reflect a majority of the same information as do textbooks from the 50’s and 60’s. In summary, Biology (sometimes referred to as life science) usually begins with an introduction to the parts and function of the microscope, followed by a brief overview of biochemistry and its relation to the living world. This then leads to a study of the functions and parts of a cell, basic genetics, and a kingdom-by-kingdom look at the various organisms on the planet. Then, the human anatomy is introduced, finalizing with ecology and plant systems. This curriculum outline pattern or an abbreviated version has been followed for the past 100 years in high school biology classes across the country.

Although not documented, but widely speculated it is believed that animal dissection was incorporated into the high school biology curriculum sometime between 1910 and 1920 as a method of teaching anatomy. That decade was known as the “golden era” of humane education. Instituted in Massachusetts in 1886, the program’s principles

---

*The Southeastern Conference on Biology Teaching in 1954 prepared a written summary statements of how their special fields could and should contribute to the improvement of biological education. The six fields were: 1. Heredity and development. 2. Evolution and paleontology. 3. Morphology. 4. Taxonomy. 5. Physiology and health. 6. Ecology and conservation.*
focused on compassion, goodwill and humanity toward all life specifically targeting children and animals at its core (Thompson & Gullone, 2003, Selby, 1995). By 1920 over 20 states in the union made human education programs compulsory; however, many of these states had no sanctions in place for non-compliance. This resulted in the uneven observance of humane education programs across the country (Antoncic, 2003).

Prior to that era, dissection was predominately attempted in colleges, particularly in medical schools. Documentation denotes as far back as the 1700s, medical schools required their students to dissect cadavers (a human corpse). When cadavers were difficult to obtain, dogs served as a substitute. It is believed that other animal dissections became integrated into the general college curriculum in order to better prepare students for medical school (Gelfand, 1972).

One of the reasons it is believed dissection became popular at a high school level was in response to the launching of Sputnik, science in the 1960s took on a new level of importance. Following the Biological Sciences Curriculum Study, a federally-funded initiative in the 1960s to create science curricula for elementary and secondary school students (Emmons, 1980), more high schools established advanced biology courses involving dissection of cats, minks, and fetal pigs, as well as an increased use of live animals. Previously, dissection of such animals was more common in college-level comparative anatomy courses (Emmons, 1980). In 1988, it was estimated that animal dissection occurred in 75-80% of pre-college level biology classes (Orlans, 1988).

Furthermore, the practice of dissection became well-established and engrained in the science curriculum. The wide spread acceptance was encouraging to science teachers who were stimulated into using the hands-on methods of teaching and learning. Science
drifted away from rote memorization of facts to now requiring students to design, hypothesize and carry out scientific experiments (Youngblut, 2001). Lab exercises now focused on nerve and muscle interaction, the reflexes of an organism, the various stages of embryo development, reproduction and moved away from simple dissection. Facing the realization, but considered an advancement, students in the study of science at the school-base level, were now required to study live organisms, kill them, and then analyze them some more through dissection activities (Youngblut, 2001). One common activity was the practice of frog pithing; (insertion of a needle into the spinal cord which paralyzes the frog in order to see the heart beating while dissecting) became incorporated into the curriculum (Russell, 1996). This vast majority of scientific work did not lead to medical advances and as a result millions of animals suffered and died in vain (Dagg, 2008, Greek & Greek, 2003).

It took approximately 15 years for the practice of dissection to fall into disfavor. Because of many complaints of cruelty to animals, a 1980 code of practice on animal in schools for science teachers was established (Russell, 1996). The purpose of this code was to set guidelines for the use of animals in high school biology classes. This code stated, "no experimental procedure shall be attempted in mammals, birds, reptiles, amphibians, or fish that shall cause the animal pain or discomfort or that interferes with its health" (Russell, 1996, p. 6). Many teachers eliminated activities such as pithing frogs and regressed to simple dissection.

**Organism Usage Statistics for Dissection**

Today, the majority of North American students (approximately 75-80%) will participate in at least one classroom hands-on animal dissection during their primary and
secondary years (Balcome, 2000; Orlands, 1993). Although the exact number of dissected organisms in high schools is difficult to calculate, most agree it falls somewhere between 5 and 6 million per year (Humane Society of the United States, 2004; Balcombe, 1996; DeRosa, 1986;). Frogs, fetal pigs and cats are the highest selected animals used in a dissection activity, but other popular vertebrate dissection specimens "include rats, dogfish sharks, perch, pigeons, salamanders, rabbits, mice, turtles, snakes, mink, foxes and bats. Invertebrate species include crayfish, grasshoppers, earthworms, clams, sea stars, squid, sea urchins and cockroaches" (Humane Society of the United States, 2004, 2) according the Humane Society of the United States (HSUS).

With the rise in animal dissection in the classroom, there has been an increase in supply laboratories that stock these organisms, and although supply laboratories claimed to raise these animals on site, the HSUS (2004) claimed that most of the dissection organisms are not raised in the laboratories at all but, rather are collected from the wild by these companies. Some of these organisms included "frogs, spiny dogfish (sharks), mudpuppies and other salamanders, birds, snakes, turtles, fish and most invertebrates"(3). Others not caught in the wild came from animal breeders, dealers, shelters, pounds, fur farms and slaughterhouses (Balcome, 2000). Without question, there is a convenience in purchasing through biological supply companies, as the animals can be purchased in bulk and delivered directly to school along with any other required lab equipment such as pans, scalpels, pins, etc. (Hart et al., 2008).

Bullfrogs, a common dissection specimen, experienced a 50% decline in population in this country between the mid-1950s and the early 1970s. Although many factors contributed to this decline, supply companies experienced difficulty obtaining
enough to meet their customers' needs. They began purchasing these amphibians from other countries, namely Mexico and Canada, to fulfill the demand (Weil, 1996).

However, other countries are starting to notice the benefits of preserving their amphibian population and not exporting them. Until recently, India was a large exporter of frogs for frog legs. It banned this export despite the fact that the country was earning ten million dollars a year exporting this product. India soon realized it was spending ten times as much importing chemical pesticides to fight insect infestations directly attributable to the frog shortage it created. In the US, a similar trend is noticeable. "The dwindling numbers of frogs in the wild have resulted in increased pesticide use, habitat destruction, and other forms of environmental deterioration." (Jackson, 1991, Environmental Groans, para. 3).

**Purpose of Dissection**

Teachers have different intentions as to the purpose for the use of dissection as an activity in their classroom. Most use it to simply provide students a general overview of the anatomy of vertebrate and invertebrate organisms. The National Science Teachers Association (NSTA) claims the "interaction with organisms is one of the most effective methods of achieving many of the goals outlined in the National Science Education Standards (NSES)" (NSTA, 2005). Laboratory activities are generally conducted to illustrate and reinforce the scientific method of thinking. They foster organizational, inquiry and higher order thinking skills.

This belief is not recent; it was reinforced approximately thirty years ago in 1975 with a report published by the Royal Society Education Committee. The Royal Society is a self-governing fellowship of many of the world’s most distinguished scientists drawn
from all areas of science, engineering, and medicine. Founded in the 1660s, the society’s mission is to recognize, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity (The Royal Society, 2012). Primarily the Society felt dissection was necessary for students to gain knowledge of internal structures of organisms. They felt there was no viable substitute for the real organism and that the knowledge was more likely to be remembered when a student partakes in the authentic activity (The Dissection of Animals in Schools, 1975). Additionally the report stated that dissection is a learning tool on many different levels. It teaches students to respect life and to help them understand that humans kill animals for a number of reasons including food, education and sport. Within the aspect of education, dissection helps students become acclimated to many professions in the medical field and therefore gives them insight to jobs that they could possibly pursue in the future. The report also notes that dissection teaches manipulative and decision making skills. It also teaches students to record and analyze their observations and data (The Dissection of Animals in Schools, 1975).

This view is reinforced by Bernstein (2000) (a leading researcher who advocates for dissection in research). She says that on the surface dissection is often viewed simply as an anatomy lesson, but in reality, it is much more. It is a way to teach skills that are essential to all science. These skills include "making observations, forming hypotheses, experimenting, analyzing results and drawing conclusions" (p. 374).

Balcombe (2001) disagrees with the usage of dissection. He noted that tradition is what drives the continuation of conducting dissection activities and it is no longer a
pedagogical necessity. He states that hands-on dissection is "weak on both concept learning and problem solving" (p. 5).

Marr (2001) agreed with Balcombe. She stated that because today's teachers were taught using a dissection driven curriculum, dissection is all they know. It therefore follows suit that they would teach in a similar fashion. She continued by noting that "repetitious memorization of terms and body parts serves little practical gain except to be something that can be tested" (p. 9). Gilmore (1991) adds to this thought by saying biology teachers have not had the necessary training in college to teach any other way than by dissection (p. 272).

The argument continues for and against dissection, one can say that it provides many benefits for those students that are contemplating a profession in the veterinary or medical fields. In a pig dissection study conducted by Barr and Herzog (2001), it was noted that the activity increased the students feeling positively or negatively for dissection. By conducting the activity, the students themselves reinforced this notion and reported that the dissection experience helped them decide for, or against, a certain career path in science. Additionally, some students felt dissection should be reserved only for students who are planning to study science in college. The students felt that by providing these students the opportunity to dissect, it would allow them to learn the rudiments of a possible future practice (Barr & Herzog, 2001).

Balcombe (2001) believed this advance preparation is unnecessary. He noted that with the advancement of technology and artificial resources, veterinary schools are using synthetic skeletons and models of animal organs to teach basic surgical procedures as well as treatment of common ailments such as the resetting of broken bones. With this
directional shift in pedagogical presentational methods, veterinary schools are reducing, not increasing the amount of dissections students are required to perform.

**Proponents for Use of Hands-On Dissection**

Despite the arguments presented by those against dissection, proponents of dissection firmly believe there is no substitution. They believe that nothing compares to the actual sight, smell, and feel of an animal's tissues and organs. It has been noted that animal dissection engages four of the five senses whereas alternatives engage two, at most. Through traditional dissection, students not only get to view the spatial relationships between organs, they can compare their sizes and densities. This, experience proponents claim, is not capable effectively in any other way.

Dissection also provides firsthand knowledge and hands-on learning. Traditional dissection allows students to make a correlation between what they have seen and read about, connecting the literature content with the practical experience. "You can have a student regurgitate on a paper-and-pencil test that a mammal's lungs are spongy, but there is no way that a student will understand what spongy means unless they see a real lung" (Offner, 1993, p. 148). It confirms what students have learned and provides a permanent, irreplaceable learning experience that cannot be duplicated.

Valli (2001) believed students need to experience working with actual animals and that they need to feel and smell fresh or fixed animal tissue. He stated that "although alternative methods of learning anatomy may be sufficient for factual regurgitation, most cannot provide the real-life, three-dimensional understanding of anatomy learned in hands-on dissection" (p. 5). McInerney (1993) agreed with this concept. He stated that "molded plastic, cotton and video display terminals provide neither the sensory nor the
informational content of actual tissues and organs" and that to portray them as a viable substitution is doing a grave injustice to the students (p. 277).

Furthermore, dissection affords students the opportunity to study variation within a species. They see that life is not an absolute, but varies from specimen to specimen. Through dissection, students are provided a clearer understanding of their own bodies and get to see them as an intricate working of cells, tissues and organs (Bernstein, 2000; McInerney, 1993). Textbook diagrams and technological images have not advanced enough to provide the realism of an actual organism. Those means of informational content can only hope to mimic the real thing. Morrison (1992) supported dissection by stating that, "dissection is simple direct science." In dissection, "the only pieces of equipment required are the student's own hands, eyes, and brain. Nothing else intervenes between the student and their observations, a rarity in modern science" (p. 22).

Moreover, with dissection, the procedure does not always go as planned. In science and life in general there is inherent chaos. There is an unspoken benefit for students to learn to expect surprises and not to assume everything will work simply because they wish it to (Russo, 1997). With a virtual activity, there is no element of surprise. Students can learn no more than these and the "emotional punch" is taken out of the dissection (Russo, 1997, p. 581). Specific outcomes are a given. Russo also believes that virtual programs allow students to "step further and further away from real life, messy and unpredictable as it is, full of emotional and ethical difficulties" (p. 581) and this is detrimental to their learning.

Bernstein (2000) brought up another benefit of dissection. She stated that in addition to basic anatomy, by comparing and contrasting the internal structure of various
organisms, students can see how organisms have changed over time. Through this learning process students gain a better understanding of the concept of evolution.

Wheeler (1993) expanded on this, stating that when using models, the organs just become abstract names or words. Through the personal experience of dissection these names become concrete and become much more meaningful and their memory will "persist for their lifetime" (p. 31).

Do these benefits justify the taking of an animal's life? Resoundingly, proponents say yes. They contend that these animals are not killed needlessly. Fetal pigs come from sows that are already slated for slaughter and there is an over abundance of animals being killed needlessly in animal shelters already. They believe some learning must occur through reality experiences. For example, no one would want to be on the road with someone who learned to drive virtually.

Morrison (1992) defended dissection in a different way. He stated that in science to kill an animal for research is justifiable and natural. It is no different from "a wolf that kills a deer" or "a frog that kills a fly" (p. 21). Furthermore, only humans grasp the concept of cruelty and therefore unrealistically worry about the welfare of other species. He continued to state that by using animals in the classroom, students learn respect for them and it is his opinion that children are not learning respect for animals at home because he claims that pets are often the most neglected and abused of animals.

Additionally, according to McInerney (1993) we, as a society, have become far removed from nature. Because only four percent of the US population now lives on farms, few children have ever seen an animal slaughtered for food. Coupled with this, television portrays a very gentle, beautiful side of nature. The majority of animal based
television shows promote the protection of species and glamorize the diversity of life. More often than not, animals are portrayed as helpless, harmless, and in need of protection by humans. Because of this portrayal, few people understand the carnage associated with nature or the brutality that exists in the natural world.

Prominent organizations have publicly declared their support for traditional dissection for a number of other reasons. The American Medical Association (AMA) claimed that without dissection there would not be a vaccine for polio, there would be no organ transplants, and current research in AIDS and Alzheimer's would come to a screeching halt (Riechard, 1993). They note dissection is necessary for continued medical progress. The military, also a strong supporter of dissection, believed that substitutions are not an adequate replacement for the actual animal dissection (Anderson, 1992). Their stance is that the practical experience is needed because the activity (using animals) gives their medical teams the best preparation scenario for treating soldiers in combat situations. Finally, NSTA (2005) felt that the choice of teaching methods should be left solely with the teacher; they also warn teachers to be wary of the limitations of alternative methods. They opposed interference by any governing body that would infringe on an educator's right to choose hands-on dissection. Offner (1993) eloquently sums up her stance on traditional dissection by stating, "the alternative to dissection is ignorance" (p. 148)

Recent studies have been conducted at all levels of education to determine whether traditional hands-on dissection produces better academic results than the alternative (virtual). A review of a few recent documented school-based research studies in which students who obtained content knowledge using hands-on dissection out
performed those using virtual dissection programs in science classrooms in the United States, are described here in chronological order. Michael-Clark (2003) performed a doctoral research study that examined the effectiveness of virtual simulation dissection (Digital Frog) versus hands-on dissection, focusing on frog anatomy content knowledge. The participants were 115 biology students in high school, taught by two teachers with four classes total. Each teacher selected one class to conduct a hands-on dissection, while the other class completed the simulation activity. Student content knowledge was measured by analyzing two post-dissection tests utilizing actual frogs: the first, two days later after both activities were completed and the second, two weeks later after the specific chapter was completed. The results were distinct. The students who participated in the hands-on dissection scored significantly higher on both post-tests than those who participated in the virtual activity. Furthermore, the post-tests results were consistent across race and gender.

A Cross and Cross (2004) study involving four AP Biology classes over two years was performed to determine the usefulness of a software program called Biology Frog Dissection and its effect on student achievement. The students were given the option to choose the method (hands-on or virtual) in order to complete the activity. After the activity the students were given a lab practicum assessment. The results disclosed that the students who undertook the hands-on dissection activity outperformed the students who used the biology software program significantly. Additionally, it was noted that the students who conducted the hands-on activity were better adept at extrapolating the information in answering virtual questions however, the opposite was not true (Cross & Cross, 2004).
A doctoral dissertation by Montgomery (2008) focused on the frog dissection program *Cyber Ed Dissection Series* and its effectiveness on student achievement as compared to traditional hands-on dissection. The 84 biology students in the study were from a southern New Jersey high school. The participants were divided into three groups: participants who performed the hands-on dissection, those who completed the virtual dissection and individuals who were given a choice. All students took a pre-test at the beginning of the unit. Upon completing their specific activity, all students were given a post-test of general knowledge questions and a lab practicum using actual frogs. The findings reported no significant difference academically between both methods, but a noteworthy difference was documented based on the results from the lab practicum. The students who performed the hands-on dissection outperformed those that completed the virtual dissection. Furthermore there were no substantial differences among sub-groups in relation to gender or ethnicity. Table 1 summarizes the empirical studies comparing hands-on dissection to virtual dissection that found the hands-on method to be more effective.
Table 1. Published Studies Noting Traditional Hands-on Dissection to be More Effective than Virtual Dissection.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Location</th>
<th>Principal Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michel-Clark</td>
<td>115 biology students</td>
<td>Nevada</td>
<td>The students who participated in the hands-on dissection scored significantly higher on both post-tests than those who participated in the virtual activity.</td>
</tr>
<tr>
<td>(2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross &amp; Cross</td>
<td>41 11th-12th grade AP</td>
<td>Georgia</td>
<td>The students who undertook the hands-on dissection activity outperformed the students who used the biology software program significantly.</td>
</tr>
<tr>
<td>(2004)</td>
<td>biology students</td>
<td></td>
<td>Additionally, it was noted that the students who conducted the hands-on activity were better adept at extrapolating the information in answering questions on the virtual program.</td>
</tr>
<tr>
<td>Montgomery</td>
<td>84 biology students</td>
<td>New Jersey</td>
<td>The findings reported no significant difference academically between both methods, but a noteworthy difference was documented based on the results from the lab practicum. The students who performed the hand-on dissection outperformed those that completed the virtual dissection.</td>
</tr>
<tr>
<td>(2008)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Opponents against Hands-On Dissection

Dissection is nothing more than a waste of an animal's life claim opponents to the activity. One key opponent Russell (1996) believes that dissection has become nothing more than a basic anatomy lesson where animals are substituted for humans. He believes most animal dissections are used to study human anatomy or physiology. Also, the learning is limited. Students do not learn about the animal’s traits, its environment, and the effect that the environment has on the animal. All they learn is the location of the organs inside the cavity. "The rat is not being investigated for its own sake but, rather as a cheap and readily accessible subject for learning about how the human body is supposed to work." (p. 8)
Balcombe (1996, 1997, 1998, 2000, 2001), perhaps the most outspoken opponent of dissection, contested the statement that there is no replacement for dissection. He considered that statement to be nothing more than "empty rhetoric." He contended further that dissection is not an experiment, it is a demonstration. There is no research nor any joy of discovery involved (Balcombe, 2001). He argues that there is nothing new to dissection that is not already known nor could be just as easily obtained from pictures, models or videos. Furthermore Balcombe stated that alternatives are at least as effective as hands-on dissection.

An added problem with hands-on dissection is the limited frequency in which the activity can be conducted using the same specimen, which in most cases would be once. During the activity specimens become destroyed in the process. Therefore students were prevented from re-examining the organs and their location through traditional dissection. However, dissection alternatives, such as models or computerized dissection simulations allow the process to be done a number of times without penalty. Students are able to repeat, replace or reverse their steps during the activity. Through repeated practice a student is able to build a mental capacity for the knowledge. Textley (1992) expanded this thought, and noted that in order to transfer facts from short-term memory to long-term, repetition is required.

Additionally, Predavec (2001) found that computer-based alternatives enable students to progress at their own pace. It also affords them the opportunity to manipulate a number of variables at the same time (Lazarowitz & Huppert, 1993). Students could see a range of animals of all sexes and stages of development. In Predavec's study it was noted that students who completed the virtual dissection activity scored higher on recall
assessments than those who actually dissected, and “not only were the students better able to identify structures in pictures and relate their functions, there were also better able to identify structures in real dissected rats” (p.78).

Furthermore, Maloney claimed that using a virtual program made it easier on the teacher. Once the instructions were given, the teachers mainly facilitated the understanding of content by circulating about the classroom, answering questions as needed. Maloney (2002) states that her students "completed the virtual dissection with very little instruction from the teachers, as opposed to the teachers of the students who actually dissected had to do much more demonstration and give many more instructions." (p. 571)

Although visual presentation of virtual programs have advanced significantly over time, proponents of traditional dissection have still stated there is no replacement for the feel, smell and texture of animal organs. They claimed that computer simulations are not realistic enough. Doctor Nancy L. Harrison MD, of the Scripps Memorial Hospital disputed that claim. She stated that “computer images of well-preserved tissues look more like the 'real thing' than the squishy gray organs of a formalin-fixed specimen." (Good Medicine, 2004, p. 11)

In contrast to statements made by Morrison, opponents such as Russell (1996) believed dissection led to indifference and uncaring attitudes in students. He believed so strongly in this ideal that he published a human physiology laboratory manual which contains experiments that are not invasive to the study subjects (p. 8). Russell’s approach to teaching science was one that focused on the wholeness of life and not just the mechanics of the dissection activity. Dedicated to how adults treat animals and how
animal treatment is presented to students shapes and determines their attitudes and future
treatment of animals. Russell compares dissection to a clock. He says that although
taking apart a clock is an interesting activity, if it is not done with purpose and passion
what you end up with are many pieces or useless parts. When life is treated in the same
way it fosters the belief that it is nothing more than something that can be disassembled
with no meaningful purpose.

Bowd (1993) agreed he believed dissection is an archaic, ineffective method of
teaching, and teachers need to not view alternatives as supplements but rather as viable
replacements to hands-on dissection. He further stated that teaching students to respect
life should be the ultimate goal of science teachers at the school-base level. Downie and
Meadows (1995) eloquently asked "can we justify the sacrifice of the lives of countless
small mammals for the sole purpose of helping biology students to learn already well-
established facts?" (Discussion, para. 1)

Prior research measuring student learning with hands-on dissection versus virtual
dissection concluded that knowledge gains tend to be equivalent, cost over time are less,
students are generally positive when using virtual programs and the virtual programs
provide better support for weaker students (Balcombe, 2003). A review of a few recent
documented school-based research studies in which students who obtained content
knowledge using virtual dissection programs outperformed those using hands-on
dissection in science classrooms in the United States, are here described in chronological
order. Youngblut's (2001) doctoral research set out to determine whether Digital Frog II
(a virtual frog dissection program) would be a viable alternative to the hands-on activity
in driving student achievement for seventh-grade students. A total of four classes
participated in the study, with two (50 students) that were assigned to perform the virtual activity and the other two (58 students) using the traditional hands-on method. Upon completion of the activities, a post-test was administered. The results showed that the students who used the virtual program outperform the hands-on group. Additionally, it was noted that there was a correlation between the simulation experience, the students’ attitude toward dissection and their achievement score. Furthermore, a major finding discovered, showed that less time (44%) was needed for the students to complete the virtual activity as opposed to the hands-on activity.

The research study conducted by Kopec (2002), investigated whether the virtual program *Net Frog* could compare to the hands-on dissection method in improving student achievement. Kopec used 218 high school biology students divided approximately evenly. The students were also compared according to ability levels (Honors, General Ability, and Foundations Level). In addition, the students were given pre and post-test to gauge their achievement levels. The results indicated that there were no significant differences in achievement between the *Net Frog* program and the hands-on method. Therefore, Kopec determined that the virtual program was a viable substitute to the hands-on approach.

Maloney’s (2005) study focused on determining whether a virtual pig dissection program could be used as a practical substitute to the hands-on method. In this study, 224 contributors were used from an all-girls high school. The girls were enrolled in biology classes taught by three different teachers. In the study, 88 students completed the hands-on dissection activity, while the remaining 136 conducted a virtual activity. Following the activities, the students were given a post-test to determine acquisitional
knowledge, and a lab practicum of organs and structures to be identified. The results pointed out that the students who completed the virtual activity significantly scored higher than the hands-on group on both tests. Maloney also noted that the results should not be generalized to girls that are enrolled in mixed-gender classrooms, and that further research could be conducted to determine if gender was a variable in the results. Table 2 summarizes the empirical studies comparing hands-on dissection to virtual dissection that found the virtual method to be more effective.

Table 2. Published Studies Noting Virtual Dissection to be More Effective than Traditional Hands-on Dissection.

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Participants</th>
<th>Location</th>
<th>Principal Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youngblut (2001)</td>
<td>108 middle school students</td>
<td>Virginia</td>
<td>The results showed that the students who used the virtual program outperform the hands-on group. Additionally, it was noted that there was a correlation between the simulation experience, the students' attitude toward dissection and their achievement score.</td>
</tr>
<tr>
<td>Kopec (2002)</td>
<td>218 biology students</td>
<td>Northeastern United States</td>
<td>The results indicated that there were no significant differences in achievement between the Net Frog program and the hands-on method. Therefore, Kopec determined that the virtual program was a viable substitute to the hands-on approach.</td>
</tr>
<tr>
<td>Maloney (2005)</td>
<td>224 all teenage female biology students</td>
<td>Louisiana</td>
<td>The results pointed out that the students who completed the virtual activity significantly scored higher than the hands-on group on both tests.</td>
</tr>
</tbody>
</table>
Health and Environment Arguments

Opponents to dissection not only voice their disapproval towards using animals for research, but the main concern for disagreement is the treatment and care of the animal itself. Opponents claim the animals slated for dissection are viewed as profit to supply companies. Jukes and Chiuia (2003) wrote that the “capture, breeding, housing, killing, preservation and transportation of millions of animals each year has a significant impact on the environment” (p. 35). Procuring animals for dissection can mean a disruption to an ecosystem given that many dissected animals are caught in the wild (Rosenberger, 1998).

Balcombe (1996) cited a 1971 article in Bioscience magazine that described how frogs were captured and treated. It documented the shipment of live frogs and how they were exposed to "dehydration, overheating, freezing, rotting, crushing and death" (Balcombe, 1996, p. 24). The article, according to Balcombe claimed that exposure to extreme temperature killed up to half of the 20,000 to 30,000 frogs waiting for shipment. Although published in 1971, Balcombe believes that very little has changed with this practice, and there is still some validity to the article.

In 1990, the People for the Ethical Treatment of Animals (PETA) acquired undercover footage of the Carolina Biological Supply Company (CBSC). It is a company that markets the sale of organisms for dissection. The video showed employees poking cats with metal hooks and the cats being stuffed into wire cages. The animals were then gassed. Although not completely dead, the cats were hooked up to formaldehyde-infusing machines and thus embalmed while still alive. Denying all claims, the company stated that the animal movements were nothing more than muscle spasms common to
most recently dead animals (Balcombe, 1998; Weil, 1996). This cruelty toward animals in laboratories and at supply companies was unknown to teachers who purchase animal carcasses before the onset of the internet. Balcombe (1998) and Weil (1996) also noted that in this same video similar treatment of rats was shown.

Chemicals

An additional problem associated with dissection is the chemical that is used during the activity. The animal carcasses used for dissection are stored in preservatives, usually in formalin (a solution that includes formaldehyde) before shipping (Jukes & Chiuia, 2003). The United States Occupational Safety and Health Administration (OSHA) classified formaldehyde as a toxic and hazardous substance. It is a known carcinogen that has been linked to nasal and lung cancers (OSHA, 2003). Students are required to wear goggles during dissection because prolonged exposure to formaldehyde can lead to watering and irritation of the eyes, sore throat, nasal discharge and discomfort, respiratory problems, burning sensations especially in broken skin (Jackson, 1991). In a qualitative study by Barr & Herzog (2001) noting student reactions to dissection, most students noted the odor of the chemicals as being a "particularly unpleasant aspect of dissection" (p. 7). The students in this study mistook the scent of decaying animal flesh and noted that although they eventually became acclimated to the smell, it was overpowering and nauseating at first, the odor was indeed caused by the chemicals (Barr & Herzog, 2001).

Recently, supply companies have begun to sell preserved specimens in solutions that are formaldehyde free. Chemicals such as Borealene II, are now used because of their low odor and non-irritating effects. The issue presented in purchasing these
formaldehyde-free animal specimens is the higher price and limited selection. Now teachers who are faced with tight budgets must now decide between animal availability, health concerns and cost.

**Diseases**

The advancement of science and technology around the world has led to the influx of animal biological supply companies entering the market. This has raised concerns of diseases the dissection specimens may carry. Countries such as China have used animals in dissection activities that have been found to carry rabies (Sharma, 2006). Although this has not been a problem in the United States, any rat used for classroom dissection must have documentation noting the country of origin, and that the animal has been quarantined and is free of rabies (Sharma, 2006).

**Cost and Safety**

States across the nation have made steep cuts to education funding since the start of the recession (see Figure 3). Because these states relied heavily on spending reductions in response to the recession, rather than on a more balanced mix of spending cuts and revenue increases, funding for schools and other public services fell sharply (Oliff, Mai, & Leachman, 2013). School districts were left searching for a way to provide a quality education, with fewer resources.

---

3 Example, as of February 18, 2013, the Boreal Northwest Biological Supply website http://boreal.com, listed a formaldehyde preserved fetal pig specimen cost at $24.45, whereas a formaldehyde free fetal pig specimen cost $36.30. A vacuum package of 10 regularly preserved rats was listed at $109.00, as opposed to a formaldehyde free package of rats costing $153.00. Also, various animals and insects are not available formaldehyde-free including grasshoppers, crayfish, mice, minks, turtles, and snakes.
At the school-based level, principals are given the autonomy to find cost saving measures that produced “at least” the same if not better academic results in the various content areas. In science, there seems to be no debate that alternatives used in a dissection activity are more cost effective than traditional dissection. Factors such as: the variety of dissected organisms (for various activities), student enrollment, and the size and complexity of the organisms make it difficult to calculate precisely how much of a classroom teacher’s budget is allocated to dissection specimens yearly. Additionally, dissection tools must be purchased, repaired or replaced periodically. To compound the

Figure 3. Percentage of Change in State Funding since 2008 (Center on Budget Policies and Priorities, 2013).
issue due to inflation, the cost of dissection specimens and equipment has continued to rise, making dissection alternatives a viable option.

Although it is challenging to come up with a precise formula to determine the cost effectiveness of virtual dissection, Balcombe (1998) had one such calculation. He calculated and analyzed the savings for teachers who had their students dissect a cat, pig and frog. Over a period of three years, he found that a teacher's savings would be over $5,000 if they switched to virtual dissection for these three organisms. This assumption was based on a teacher having no materials or equipment to begin a dissection activity, the number of organisms used equated to one per every two students, and Balcombe accounted for the price of the software itself, no hardware.

While the savings associated with virtual dissection will vary from teacher to teacher, the safety regarding computer dissections and virtual dissections cannot be debated. Virtual dissection is far safer than the hands-on method. Students do not come in contact with sharp objects in which they can accidentally cut themselves. Teachers and students do not have to worry about appropriately discarding dead carcasses in order to meet safety standards. Both parties do not have to address ingesting or inhaling harmful chemicals that might cause negative reactions. Finally, the class does not have burden of making sure that the chemicals used are suitably disposed of not affecting the environment.

**Summary**

Based on review of literatures related to the research topic, several conclusions may be drawn.
1. It is evident that researchers, educators and animal activist, have continued to debate the impact of dissection activities in helping students’ gain the foundational knowledge in the subject area and an appreciation for science.

2. It is also evident that the hands-on dissection activity as practiced in high school science classrooms has drawn disapproval from animal activists who claimed that it was shallow in scope, placed too much emphasis on the procedure and less on the organism and its impact on the ecology web, which they argued have compromised the impact of laboratory activities on students’ learning of science.

3. The need to improve student science achievement scores on national and international assessments, have pre-occupied science educators and researchers for decades.

4. It is clear from the review of the literature that the potential impact of virtual dissection programs in science education has recharged the debate on how to provide meaningful laboratory experiences for grade-school students’ in helping them understand objectives biology and science related courses.

5. A review of literature showed that several researchers have conducted meta-analyses when comparing the effectiveness of virtual dissection programs against the traditional hands-on activity to compare the effectiveness of instruction in raising students’ science outcomes similar to the study that will be presented in this report.
CHAPTER 3: METHODOLOGY

This chapter presents the methodology that was used in this study to assess the impact of a virtual dissection activity versus the traditional hands-on method on student achievement. It begins with a restatement of the purpose and research questions addressed by the study. Then an overview of the study’s design given, followed by a description of the setting, sampling techniques, procedures, description of the student activity being assessed, the instrumentation and ethical considerations.

**Purpose and Research Questions**

The purpose of this study was to investigate the impact of a student activity designed to improve science content knowledge. This student activity helped determine whether a virtual dissection program called Rat Dissection 1.1, created by Emantras Inc., offers students an effective alternative to traditional rat dissection, by allowing them to score at least as well on the same assessment. The first two research questions specifically addressed the potential impact of the virtual student activity.

1. What effect do virtual animal dissection and traditional dissection methods have on student achievement as measured by a general animal (rat) anatomy assessment (consisting of recall questions), and a lab practicum exercise (consisting of organ identification and function)?

2. What effect does virtual dissection (as opposed to the traditional method) have on student achievement differences, based on gender and the primary ethnicities (Caucasian, African-American, and Hispanic)?

The final research question was designed to gather information from the students that will use both methods (virtual and traditional) in implementing the activity. In order to assess
the impact of virtual dissection, it is important to determine the students' preference to both methods. This data was collected through the administration of a qualitative student survey:

3. What preference do high school biology students have as a means for acquiring the knowledge of human systems: virtual dissection of organisms or the traditional dissection method?

Setting

This study was conducted in a large (over 2,500 students) high school in a district located in Central-South Florida. The school was highly diverse with a significant number of students (70%) listed as receiving free or reduced lunch. It was labeled a Title I school. This categorization was based on the students' low socioeconomic status, which distinguishes the school as being eligible for federal financial resources to aid in educating those students.

Comparative schools within, and in nearby school districts are measured against the high school where the study took place. These schools consistently earn higher grades from the state on the accountability grading system. Nevertheless, in recent years, the school made great strides in improving student achievement. Innovative leadership and proactive implementation of ideas into programs resulted in data-driven, reflective teaching that drives the learning environment. However, science scores remained lower than the state average and suggest a need for specific student engaging methods to acquire the information.
Target Population

The target population was high school students enrolled in a Biology course. Biology is offered in the 10th grade. Due to the fact that, the district uses a controlled-choice model within carefully selected zones, the student population at the selected high school was fairly representative of the St. Lucie County School District as a whole. The respective high school has struggled to achieve high scores on the state end of course biology assessment making it an ideal candidate for this study. Therefore, although this study examined just one high school, it was expected that results could be generalizable to other similar high schools within and outside the district.

Sampling Plan

All of the biology instructors (5) in the science department were presented with the idea of the research. Three out of the five teachers elected to participate. The final two elected not to participate for various reasons that were respected. The teachers that elected to participate have volunteered all of their classes – twenty in total (one teacher instructs a class in Environmental Science). The sample size was approximately 325 students, encompassing roughly 60 percent of students enrolled in a biology course at the respective school.

For this study, there were two aspects considered in order to categorize a student as a participant in this study. First, a student had to have passed their previous science course in the ninth grade. Prior content knowledge is the foundation upon which new information is comprehended and learned. “A learners existing knowledge has a large impact on knowledge acquisition (O’Reilly & McNamara, 2007, p. 163).” Furthermore, according to Pressley, El-Dinary, Marks, Brown, and Stein, (1992), new information that
is easily integrated with existing knowledge is remembered better. Any student that did not pass the previous science course in the ninth grade was excluded from this study.

Second, in order for a student’s data to be accepted for this study, the individual had to be in compliance according to the state’s requirements for full-time equivalent (FTE) student status. FTE was defined as one student in membership in a program or a group of programs for 900 hours (25 hours per week) for a 180-day school year. The survey occurred twice during each 180-day school year -- once in October and once in February (Stlucieschools.org, 2012). A student was eligible for the study if the student met both of the following conditions for FTE acceptance in October 2012:

1. Student was in program membership at least one day during the survey week in a St. Lucie County public high school or alternative high school, and
2. Student was in attendance at least one of the days of the survey period or one of the six days preceding the survey period for which attendance was scheduled (Stlucieschools.org, 2012).

This determination was important because it created a foundation for eliminating any students from the study that transfer to the school during the 2012-2013 academic year. It was the intent of the researcher to give the teachers of record, as much time presenting prior content material in preparation for the study.

Research Design

This research was designed to assess the impact of a virtual rat dissection program versus the traditional hands-on dissection method on student achievement as measured by teacher created pre/post assessments, and a lab practicum exercise (see Appendix A, B, & C), as well as through the measurement of student preferences (see Appendix D). Rat
Dissection 1.1 is a virtual technology program used to improve science content knowledge, while eliminating the look, feel and smell of a traditional dissection. At the time of the research, no teachers at the school where the study took place used Rat Dissection 1.1 as a part of their classroom instruction/activities.

The mixed-methods design included two components, a quantitative component in which two groups were given a pre-and post-test. Both groups were assessed with a pre-test (dependent variable). Next, one group received the experimental treatment (virtual dissection), while the other did not (traditional dissection). Then, both groups were post-tested (second dependent variable), with the results to be analyzed afterwards. The analysis was on the impact of the virtual rat dissection program on teacher-created pre/post assessments, and the lab practicum exercise. The second component was qualitative and explored the students' preference in method of knowledge acquisition of human systems of the students' in survey form.

The focus of the study was on the impact of the two methods of dissection (the independent variable), on student achievement, and student preference, (the dependent variables). Student achievement was measured using teacher-created pre/post assessments, and a lab practicum exercise. The data was analyzed by comparing means using t-test, and Analyses of Variance tests. All analyses were carried through using SPSS 20.0 for Windows Student Version with an alpha level of 0.05.

This study included one independent variable: the method of rat dissection (traditional or virtual) in a student activity. Using factorial MANOVA analysis permitted the analysis of the effect of the treatment on multiple dependent variables. In this study, the variables were all based on biology scores. Students were given a pre-test and two
subsequent tests in biology (human systems) content leading to a repeated-measures design. Factorial MANOVA identified effects across the multiple tests. Additionally, factorial MANOVA allowed for the examination of the interaction of other variables that might influence science scores other than the treatment such as, gender, and ethnicity.

**Procedures**

The student needed to submit a student/parent consent form signed and dated. This form was an acknowledgement that the parent/guardian was aware that their student was an active participant in a research study approved by the school district (see Appendix E). It also clearly stated the purpose of this study, which was to see if there was a difference in how well students learned rat anatomy when using a virtual rat dissection program versus the traditional hands-on method. The researcher wanted to determine if there was a difference in achievement and preferred method between students who dissected a real rat and those who completed the dissection virtually. To reassure that there were no known risks involved, only those students who signed and had a parent/guardian signature on the consent/assent form was able to participate in the study and have their data used. The student’s grade was not affected in anyway by choosing to participate or refusing to participate in the study. Additionally, the confidentiality of records and the identity of the student were and will be protected to the extent allowed by law. No student was personally identified in any reports or publications that result from this study.

The use of Rat Dissection 1.1 in this pre-test/post-test control group design was determined through a collaborative effort between the assistant principal that oversaw the science department, and biology teachers’ familiar with biology-based student activities.
All parties were in agreement for the use of the software. Teacher-created assessments and a lab practicum exercise for students using virtual dissection defined the experimental group. Teacher-created assessments and a lab practicum exercise of students who did not use virtual dissection, was the control group. The pretest/posttest control group design was selected because the student assignment would be determined within each teacher’s class. A fully experimental study would randomly assign the virtual dissection activity to the students in an effort to eliminate the teacher selection effect on the outcome. Often, in the case of educational research, it can be a challenge to control all factors in an investigative study due to institutional and cultural factors.

All teachers of biology are expected to align their instruction with the district’s scope and sequence for the course. The scope and sequence identifies which benchmarks from the Next Generation Sunshine State Standards in science should be taught, and when within each quarter. Initially, there is a pre-test at the beginning of the school year. As the year progresses, the district assesses student learning of the benchmarks at the end of each quarter. The third quarter district benchmark assessment could not be considered a post-test, since it was not cumulative of all standards to date on the scope and sequence. It only assessed students on the benchmark or strands for that specific quarter.

The Rat Dissection 1.1 virtual dissection program afforded students the experience of dissecting a rat, while eliminating the effect of having to handle the real animal cadaver. This technological activity allows students to acquire the knowledge needed on key standards involving human systems that are foundational to the study and understanding of biology. In the third quarter of the fiscal school year, twenty-two standards are identified in the scope and sequence for biology. Human systems address eight of those standards in the quarter. In addition to addressing specific science content,
the Rat Dissection 1.1 program provides the students the ability to “go back through the lab over and over again. This gives students the opportunity to review learning and process, something that isn’t possible with real rat dissections” (IPAD Curriculum, 2011, par. 2).

The Rat Dissection 1.1 virtual program allows the student the opportunity to perform a rat dissection utilizing the same instruments that are used in the traditional hands-on method. As the students conduct the lab activity, the program presents a simulation of the traditional hands-on lab dissection procedures. Using a touch screen method, the students followed the voice enhanced step-by-step process to complete the activity. Towards the end of the activity, the students are able to select each organ individually to which they then could view the organ as a 3-D object, rotate the organ and learn about its function within the body. Finally, the program comes equipped with an interactive quiz in which the students can assess themselves for knowledge and understanding.

Prior to covering the topic on human systems, the students are administered a pre-test on the content. The teachers created a 20-question multiple-choice and fill-in the blank assessment, which queried the students for their general understanding of the human body’s internal organs and their function. This test was used to assess where the classes were as a whole and it gave the teachers an idea of what needed to be covered, based on the classes’ prior knowledge of the content.

For this study, the students in each participating biology class were given the opportunity to perform both methods of dissection. The determining factor as to which method a student completed first was decided by the teacher. The teacher attempted to divide the class equally in half, with one group performing the dissection using the
traditional hands-on method, in teams of two, while the second group performed their
dissection activity individually through the use of the virtual application using an iPad.

Once both activities were completed, the students were given a post-test on the
content and a lab practicum exercise. The teachers created a 25-question multiple –
choice, fill-in-the-blank, and label-the-diagram assessment, which measured the students
for their conceptual understanding of the human body’s internal organs and their
function. The lab practicum queried the students for their general knowledge on the
location of the internal organs and their function. Upon completion of the assessment
and exercise, the two groups of students then switched and used the opposite method to
perform the dissection activity once again. When finished, all students were administered
a four question choice-preference survey to inquire about choice of method. It must be
noted that any student who chose not to perform the traditional hands-on dissection
method was allowed to perform the activity using the virtual method, and subsequently
was given assignments out of the textbook to be graded during downtime. The
assignments included activities and questions in the summary sections of the various
chapters of the textbook associated with the Human Systems topic in biology. Those
students were not factored in the survey section of the study.

The researcher obtained student demographic data as well, including information
on gender, ethnicity, socioeconomic status, in addition to scores on the teacher-created
assessments, and lab practicum scores for students in classrooms using virtual dissection
and for students in classrooms using the traditional hands-on method.
Instrumentation

Instrumentation consisted of a pre-test and post-test encompassing general knowledge, as well as a lab practicum test. Both the pre-test and post-test were similar but not identical. No lab practicum pre-test was created because there was concern that the students would have an adverse reaction to the sight of a lifeless animal with a structure similar to that of a human. It was agreed that the students needed to become acclimated through their own dissections.

By the time the students got to the rat dissection, they would have made progression through the animal kingdom section on the scope and sequence for tenth grade biology culminating with the topic on human systems. The students would have dissected an owl pellet, earthworm, and bull frog. Although the dissection process is similar in exploring these three entities, the similarity pales in comparison to that of a rat, whose organ arrangement and proportion resemble that of a human. In fact, the main reason rat dissection is undertaken, is to afford the students the opportunity to see a close resemblance to the organ structure from their own bodies (Maloney, 2002).

Pre-test. Before beginning the unit on human systems, the students were administered a pretest. The teachers of record designed the pre-test using the general knowledge post-test as a guide. The test sought to establish the students’ prior knowledge of mammals. It also helped to determine their knowledge of organs and their function within the various human systems.

The pre-test consisted of 20 questions. The questions were multiple-choice in nature and tried to retrieve general knowledge information pertaining to a rat and its internal systems. These questions were a mixture of identical and different inquiries to
that on the post-test. This was to prevent students from memorizing or copying down the questions, and answers to be shared or for later use, thus leading to potential contamination of the results of the study.

**Post-test.** Upon completion of the dissection activities both groups were administered a general knowledge post-test. The post-test, created by the teachers of record, covered information that was addressed in the lectures, classwork, homework and the rat dissection activity. It consisted of 25-question multiple-choice, fill-in-the-blank, and label-the-diagram assessment. These questions were a mixture of identical and different inquiries to that on the pre-test.

**Lab Practicum Test.** The lab practicum section consisted of organ identification and organ function questions for six organs in the rat system. The students were instructed one at a time to begin. They viewed three different pre-dissected rats. Each rat was labeled differently, denoting specific structures with a dissecting pin that had a number attached to it. The numbers coincided with the numbers on the students' answer sheets. Below each number that had a blank line, there was a second line which was used for the student to list the function of the indicated organ.

**Analysis of Data**

This was a mixed-methodology study with a pre-test/post-test assessment, lab practicum and a post activity survey on the student's preference towards both methods. Along with previous science grades for each student, demographical information was recorded. The primary consideration for analyzing data using hypotheses testing procedures was the assumption that all group variances were equal or that the samples came from normal populations with the same variances. Levene's Test for Equality of
Variance is a nonparametric test that was used to test equality of variance when there are several (more than two) samples.

Statistical methods were employed in this study to explore the data and answer specific research questions. For research question 1, a one-way MANOVA was used to look at the intervention and its effect on one dichotomous independent variable against the two dependent variables. For research question 2, factorial MANOVA (2x2) was used to look at the intervention and effect on multiple nominal independent variables against the two dependent variables. Lastly, for research question 3, an open-ended survey created by the researcher (see Appendix) was used to generate information regarding preference of method from the sample of individuals. The Statistical Package for the Social Sciences (SPSS 20.0) was utilized for all statistical analyses. The use of each particular test was dependent on both the statistical assumptions and the establishment of group equivalency. A .05 level of confidence was employed in all statistical tests utilized to analyze the data.

**Threats to Internal Validity**

If conducted well, a mixed-methods research provides strong evidence for a cause and effect relationship between the independent and dependent variables. The strength of the research examines the effect one variable has on another and controls for impending explanations that might present a difference. If the research is done repeatedly, the findings can be generalized to a larger population.

However, with this type of research, a number of inherent threats rise to the forefront. Campbell and Stanley (1963), note that internal threats to this specific study...
include: history, maturation, testing instrumentation, statistical regression and experimental mortality.

**History.** For this study, history, defined as "specific events occurring between the first and second measurement in addition to the experimental variable" (Campbell & Stanley, 1963, p.5) will be a threat to internal validity. Due to the span of time between the pre-test and post-test, unanticipated factors can play a role in the student learning of the material. For this reason, and to aid in minimizing this threat, the amount of days between the administrations of the pre-test and post-test was held constant for both groups.

Additionally, any student absent on the day that the posttest was to be administered was given a different make-up assessment. The practice of giving a different, make-up test is not solely for this study. This is common practice that has been in place for the entire year. The rationale behind the thought process of giving a different assessment is that if the same test is given, it would be relatively easy for an absent student to obtain ideas related to the questions and possibly the answers themselves from someone who has taken the test previously.

**Maturation.** The high school in which the research was conducted utilizes a day 1/day 2 block scheduling format. Students take half of their classes each day, but each class is double the amount of time. For example, a typical student has seven classes on their schedule and each class meets for 90 minutes, except for their homeroom class which is fifth period which meets every day for 50 minutes. A student therefore will take 1st, 3rd, 5th and 7th on an odd day, and 2nd, 4th, 5th, and 6th on an even day. The days alternate throughout the weeks during the school year. Due to this fact, the instructors
felt that it would be better to perform the study after the Biology End of Course Exam in early May. This would give teachers as much time as possible to try to cover all the material in preparation for the exam. Although there is no mathematical formula to determine the exact amount of influence maturation has on a study, including this study, the fact that the students will have already covered the topic on human systems by the time this study is to be conducted has to be stated.

**Testing.** A test effect could have been present during this study. However because a pretest/posttest design was employed in this study, the influence of the pre-test on the post-test exam was constant for both groups. It was therefore expected that since the students were essentially given a preview of the post-test by way of the pre-test, the results would be positively influenced and inflate all post-test scores.

**Instrumentation.** Instrumentation in this study should not be a factor that could be contaminated. To ensure the validity of the exams, the teachers who participated in the study taught the course in the past, and created the assessments. Both the pre-test and post-test have specific correct or incorrect answers. There are no subjective questions to be asked. Similarly formatted tests have been administered during the current school year and the only issue encountered has been a student’s misunderstanding of a question. That was not the case for this particular study.

**Statistical Regression.** Statistical regression can be considered a threat to validity when used in an experimental study. However, due to the fact that this study involved a heterogeneous group of students and students were not selected based on any other criteria, statistical regression did not pose a risk to this study. When comparing pretest to posttest scores, the regression toward the mean was minimized.
Mortality. In addressing mortality, it was unlikely that a participant would drop out of the study between the pretest and posttest. This is due to two reasons: (1) because the research was conducted late in the school year, the chances of a student dropping out or transferring to another school minimized; and (2) because there was a short time span between the pretest and posttest (three 90 minute class days), it is unlikely that a participant would drop put between pre-test and post-test administrations. However, in the unlikely event a student misses any part of the study, the student will be given an alternative activity, as is standard practice. That student will not be considered as part of this research study.

Hawthorne Effect. The Hawthorne effect is a validation of research that places the researcher in a detached position from his/her “subjects” thereby avoiding any change of corrupted data through social engagement (Coombs & Smith, 2003). By having the research study performed after the Biology End-of-Course exam, was the belief that the pressure will be minimized for the teachers and students to conduct the activity and participate in the assessments with minimal influence by the researcher.

Threats to External Validity

Utilizing the pretest-posttest format, by nature, certain external threats are characteristic to this type of study. External threats to validity are those that limit how a study could be generalized to a larger population. These threats include the interactive effects of testing and selection biases, as well as the multiple treatment interference (Campbell & Stanley, 1963).

Interactive Effects of Testing. Although slightly different tests were used for the pretest-posttest assesments, the pretest may give the students a preview of what to
expect on the general knowledge posttest. In this particular study, because the students were pretested, they may have responded differently to the posttest. This may have made the treatment different than if the students had not been pre-tested.

**Interactive Effects of Selection Biases.** In this particular research complete randomization was impossible, in part because the students selected for the study had already been assigned to a teacher’s class section by the counselors of the school. Secondly, the teachers of record decided who began with what activity (virtual or traditional hands-on) first. Therefore, the results can only be generalized to subjects exposed to an identical selection process.

**Multiple Treatment Interference.** The subjects in this study were exposed to the treatments (activities) specifically for this study. Therefore the findings could only be generalized to individuals exposed to the same treatments in the same order of presentation.

**Ethical Considerations**

All identifying student information was erased after retrieval from the databases, (i.e., Skyward and Performance Matters) maintained by the school district under study. Teacher names will be replaced with identification numbers. The purpose of this study was not to evaluate teacher quality, but rather to determine whether virtual dissection is an effective content acquiring student activity that should be expanded across more classrooms within the school and possibly the district. Therefore, while it was necessary to examine student data by teacher to control for teacher effect, it was not necessary to include teacher names.
Summary

The researcher followed the series of steps outlined in order below:

1. Approximately 325 high school students were recruited from three biology teachers’ classes at a Title I school in the St. Lucie County School District.

2. The high school biology teachers administered a pre-test on the anatomy of a rat specimen.

3. The biology students were divided in half (based on alphabetical order), then performed the dissection activities on a rat specimen (one group conducted the traditional hands-on method and the second group performed the activity using the virtual computer program).

4. The high school biology teachers administered a post-test on the anatomy of a rat specimen.

5. The biology student groups then switched and performed the opposite dissection activity.

6. The high school biology teachers administered a post-activity survey to determine student preference for the two activities.

7. The researcher completed quantitative analysis of the pre and post-tests.

8. The researcher completed qualitative analysis of the surveys.
CHAPTER 4: DATA ANALYSIS

Descriptive Statistics

The full sample included 311 students. However, information was missing on some of the students as to which type of instruction they received: traditional or virtual. There were 301 students included with information about type of instruction: the breakdown is in Figure 4. The breakdown by gender is shown in Figure 5 and breakdown by race is shown in Figure 6.

Figure 4. Type of instruction received by students in the sample.

Figure 5. Gender breakdown of sample. Note the sample is close to even, with slightly more male students than female students.
Students in the sample were given one of two types of instruction: traditional classroom instruction with actual rat dissection or a virtual rat dissection. Students were given a pretest, a posttest, and a practicum. The posttest and practicum were administered after the rat dissection intervention. The pretest consisted of 20 questions. The pretest score used in the research was a percentage; the raw score was divided by 20 and multiplied by 100. The posttest consisted of 25 questions and was also scored as a percentage. The practicum consisted of 3 questions and was likewise scored as a percentage. Hence, all three scores are directly comparable to each other.

The mean scores for all students are shown in Figure 7. Note that the mean test score increased from the pretest (38.68) to the posttest (53.23) for all students; the practicum had only three questions and most students scored correctly on approximately two out of the three questions (69.14).
Next, Figure 8 shows the means scores for all tests administered before or after traditional dissection or virtual dissection. Again, the mean score for all students increased from the pretest to the posttest. The students that completed the hands-on traditional activity first generated a pre-test mean score of 39.79. After completing the activity, their posttest result mean score increased to 56.83, a 42.5% increase. Subsequently, the pre-test results of the students who performed the virtual activity produced a mean score of 40.50. After completing their activity, the post-test mean score increased to 45.96, a 13.5% increase.
Figure 8. Pretest, posttest, and practicum scores for all students under both the traditional and virtual dissection methods.

Figure 9 shows the breakdown of the means of all three tests by gender, where mean scores of both genders increased from the pretest to the posttest. For males, the posttest score resulted in a 42% increase from a 37.53 mean score to a 53.22 mean score. For females, the posttest score resulted in a 34.5% increase from a 39.43 mean score to a 53.07 mean score. The data also denoted males to have higher posttest scores, but lower practicum scores (65.59 to 71.85), than females.
Mean Pretest, Posttest, and Practicum Scores By Gender

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest Mean</td>
<td>37.53 (N=89)</td>
<td>53.22 (N=93)</td>
</tr>
<tr>
<td>Posttest Mean</td>
<td>53.07 (N=75)</td>
<td>65.59 (N=90)</td>
</tr>
<tr>
<td>Practicum Score</td>
<td>71.85 (N=90)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9. Pretest, posttest, and practicum scores for all students broken down by gender.

Figure 10 shows the breakdown by ethnicity. After analyzing the results, the data revealed that African-American students had the highest averaged gains from pre-test to posttest. The 48.25% increase (pre-test mean score of 33.88 to a posttest mean score of 50.26) was compounded with the highest overall practicum mean score of 71.11 for all sub-groups. The sub-group Other was next in percentage gain with a 36.38% increase, beginning with a pre-test mean score of 38.54 and ending with a 52.55 post-test mean score. The Hispanic sub-group results denoted an increase of 13.15 points, which amounted to a 33% mean increase from 39.81 pre-test score to 52.96 post-test score. The Caucasian sub-group results showed the least amount of increase from pre-test to posttest, but the group had the highest overall scores on both test. The data denoted a 43.70 pre-test mean score and ended with a 56.74 post-test mean score which resulted in a 29.75% gain.
The scores are further broken down by gender and ethnicity, and then by type of instruction received. Figure 11 shows the breakdown by type of instruction for male students only. After analyzing the results, the data revealed, male students that performed the traditional hands-on dissection showed a mean increase of 57.58%, escalating from a pre-test 34.54 mean average, to a 56.00 posttest mean average. In comparison, the male students that performed the virtual dissection activity produced a pre-test mean score of 40.91, and increased to 48.63 on the posttest assessment only resulting in an 18.88% increase.
Mean Pretest, Posttest, and Practicum Scores Under Traditional and Virtual Instruction, Males Only

Figure 11. Pretest, posttest, and practicum scores for male students only, broken down by type of instruction.

Figure 12 shows the same breakdown for female students only. Scores for female students in the traditional dissection group show a greater increase from the pretest to the posttest than females in the virtual dissection group. After analyzing the results, the data revealed, female students that performed the traditional hands-on dissection showed a mean increase of 43.88%, escalating from a pre-test 39.95 mean average, to a posttest mean average of 57.48. Comparing, the female students that performed the virtual dissection activity, the results produced a pre-test mean score of 39.6, and increased to only 41.71 on the posttest assessment resulting in a 5.33% increase.
Figure 12. Pretest, posttest, and practicum scores for female students only, broken down by type of instruction.

The next figures show the breakdown by type of instruction for students of different ethnicities. Figure 13 shows the breakdown by type of instruction for Caucasian students, Figure 14 shows the breakdown for African-American students, and finally Figure 15 shows the breakdown by type of instruction for Hispanic students. The scores show an increase for those who received traditional instruction over those that received virtual instruction for all three races.
Figure 13. Pretest, posttest, and practicum scores for Caucasian students only, broken down by type of instruction.

Figure 14. Pretest, posttest, and practicum scores for African-American students only, broken down by type of instruction.
Research Question 1

The first research question was: What effect on student achievement do virtual animal dissection and traditional dissection methods have on a general animal anatomy assessment (consisting of recall questions), and a lab practicum exercise (consisting of organ identification and function)? This question was examined using an independent samples \( t \) test and also using a MANOVA model. The independent samples \( t \) test was used to compare the two methods of dissection instruction on the difference between pretest and posttest scores. The MANOVA model permitted the analysis of the effect of the treatment on multiple dependent variables. In this study, the variables were biology scores. Students were given a pre-test and two subsequent tests in biology (human systems) content leading to a repeated-measures design. For research question 1, a
simple MANOVA model was conducted with two dependent variables, but only one independent variable: type of dissection instruction received by the student.

**Independent samples t test.** In order to run the independent samples *t* test, a new variable was created. This variable was named Score Improvement and was calculated as the difference between the pretest score and the posttest score. Figure 16 shows the average mean of score improvement for all students (15.22) and a breakdown of score improvement by type of instruction. Note that scores rose between the pretest and the posttest for all students, and for each instruction group. Posttest scores did increase much more, however, for the traditional dissection group (19.45), than the virtual dissection group (6.18). Whether this difference was significant or not was then determined by means of an independent samples *t* test.

![Mean Score Improvement, All Students Overall and By Type of Instruction](image)

Figure 16. Mean score improvement for all students and by type of instruction.

One assumption of an independent sample *t* test is that the variances of the two samples are equal. Although the *t* test can be run on data where the variances are not equal, the methodology varies somewhat. Therefore it was first necessary to test whether
the variance for the score improvement variable under traditional dissection instruction was equal to the variance for the score improvement variable under virtual dissection instruction. Levene’s test for equality of variances was used, where the null hypothesis for, was that the variances between the two groups were the same. In this case the null hypothesis was not rejected for this data at a significance level of $\alpha = .05$, with $F(158) = 0.01, p=.91$.

The null hypothesis for this $t$ test was that the mean score improvement was the same for both the traditional dissection and the virtual dissection groups. Using the $t$ test for equal variances yielded the result: $t(158) = 3.25, p = .001$. Hence the null hypothesis was rejected. The boxplot in Figure 17 shows the difference in the means of the two dissection groups. The mean score improvement was higher for the traditional dissection group; the $t$-test confirmed that this difference was significant.

![Boxplot](image)

Figure 17. Boxplot of score improvement variable for both traditional and virtual dissection groups, where t test confirmed difference is significant.
**MANOVA model.** A MANOVA model is used to examine the influences of an independent variable on more than one dependent variable jointly. For RQ1, the two dependent variables were the posttest and the practicum. The independent variable was the type of instruction, which had the two options: traditional dissection instruction and virtual dissection instruction. The assumptions of MANOVA analysis include:

1. The dependent variables are correlated, but not highly. Horn (2013) suggests that if the correlation were .60 or above, either making a composite variable (in which the highly correlated variables were summed or averaged) or eliminating one of the dependent variables should be considered as an alternative to MANOVA.

2. Homogeneity of covariance matrices, which is tested by Box’s M test.

3. Homogeneity of variance, which is tested by Levene’s test of equal variance.

4. Multivariate normality, which can be checked by examining the univariate normality for each dependent variable by means of a Q-Q plot of the dependent variables. (Bian, 2013)

Figure 18 and Figure 19 show the Q-Q plots of posttest scores and practicum scores. If normally distributed, the straight line in the plot represents the expected values (Bian, 2013). The distribution of posttest scores was approximately normal except at both ends of the distribution. The distribution of pretest scores was also approximately normal, but not continuous because of the nature of the practicum scores. A problem could exist with the use of MANOVA analysis because of the non-continuous nature of the practicum.
variable. The analysis was continued, but possible problems with the methodology in this case are discussed in Chapter 5.

Figure 18. Q-Q plot of posttest scores. Note that the distribution appears approximately normal except at both ends of the distribution.

Figure 19. Q-Q plot of practicum scores. Note that the distribution appears approximately normal, but not continuous because of the nature of the practicum scores.
A Pearson’s correlation test between the two dependent variables yielded the results, $r(166) = 0.23, p = .003$. This indicated that there was correlation between the two dependent variables, which was significant at a 1% level, but the correlation coefficient was below 0.60, thus indicating that assumption 1 listed above is not violated. The Box’s $M$ test of homogeneity of covariance matrices found that the null hypothesis of equal covariance matrices could not be rejected (Box’s $M = 3.50, F = 1.15, p = .33$). Finally, homogeneity of variance was established for both dependent variables, as the null hypothesis of equal variance for both dependent variables ($F(1,164) = 0.12, p = .73$ for posttest and $F(1,164) = 0.002, p = .97$ for practicum) could not be rejected at a 5% level.

Because of the non-significance of Box’s $M$ indicating homogeneity of covariance matrices, Wilk’s Lambda was the appropriate test to use to examine the multivariate effect of type of instruction on the posttest and practicum scores. Using an alpha level of .05, this test was significant, Wilk’s $\Lambda = 0.94, F(2, 163) = 5.55, p = .005$, partial $\eta^2 = .06$. The significant $F$ value indicated that the scores of the posttest and practicum combined were significantly dependent on which type of instruction the students received. The multivariate $\eta^2$ indicated that approximately 6% of the variance in the dependent variables was associated with type of instruction (Horn, 2013).

In order to determine how the dependent variables differed for the independent variable, the next step was to examine the univariate ANOVAs of the variables, or the tests of between-subjects effects in SPSS (Laerd Statistics, 2013). According to the ANOVA table, the type of dissection instruction had a statistically significant effect on the posttest score ($F(1, 164) = 11.02; p = .001$; partial $\eta^2 = .06$), but not on the practicum score ($F(1, 164) = 0.16; p = .69$; partial $\eta^2 = .001$). It is important to note that, in order to
account for multiple ANOVAs being run, an alpha correction, such as a Bonferroni correction, should be made. With this correction a statistical significance of $p < .025$ would be considered significant; the posttest variable met this criterion (Laerd Statistics, 2013). In conclusion, overall the MANOVA test indicated that the type of dissection instruction received by students influenced the combined grades of the posttest and the practicum. However, when examined separately, it appeared that the posttest score was influenced by the type of instruction, but the practicum score was not.

**Research Question 2**

Research question 2 was: What effect does virtual dissection (as opposed to the traditional method) have on student achievement differences, focusing on gender and the primary ethnicities (Caucasian, African-American, and Hispanic)? As with RQ1, RQ2 was examined first by means of independent samples $t$ tests and then by means of a factorial MANOVA test. This research question is answered by first examining the differences based on gender, using both methodologies, and then the differences based on ethnicity.

**Independent samples $t$ tests, differences by gender.** In order to begin to answer this research question, an independent samples $t$ test was run examining males and females separately for the virtual dissection group only. In other words, do males or females benefit the most from the virtual dissection technique, based on their score improvement? Figure 20 shows the mean for males (8.41) and females (3.25) students in the virtual dissection group. The scores improved for both genders, but increased more for male students than for female students. Whether this difference was significant or not was determined by means of an independent samples $t$ test.
Figure 20. Mean score improvement for virtual dissection students only, broken down by gender. Note that the mean score improvement appears significantly higher for male students.

As described previously, Levene’s test for equality of variances was used to test the associated assumption of an independent sample t test. In this case, the null hypothesis for Levene’s test (that the variance for male students was equal to the variance for female students) was not rejected for this data at a significance level of $\alpha = .05$, with $F(47) = 3.04, p = .09$. This meant that the $t$ value for equal variances was used in this case.

The null hypothesis for this $t$ test was that for students in the virtual dissection group only, the mean score improvement for male students was the same as the mean score improvement for female students. Using the $t$ test for equal variances yielded the result: $t(47) = 0.71, p = .48$. Hence the null hypothesis that the mean score improvement for male students was the same as the mean score improvement for female students in the virtual dissection group was not rejected. Although the score improvement for the virtual
dissection group appeared higher for males than females, the t test results illustrated that the difference was not significant at a 5% level. The boxplot in Figure 21 shows the difference in the mean score improvement by gender for the virtual dissection group only.

![Boxplot of score improvement variable for both male and female students in the virtual dissection group only.](image)

Figure 21. Boxplot of score improvement variable for both male and female students in the virtual dissection group only.

**MANOVA analysis, differences by gender.** For this question, factorial MANOVA allowed for the examination of the interaction of other variables that might influence science scores other than the type of dissection instruction, namely gender, and ethnicity. For RQ2, two different factorial MANOVA models with more than one independent variable each were employed to answer the question. Type of instruction was included as an independent variable in both models, while gender was included as an independent variable in the first model and ethnicity as an independent variable in the second model. The MANOVA model including type of instruction and gender as independent variables is explored in this subsection.
The Box’s $M$ test of homogeneity of covariance matrices found that the null hypothesis of equal covariance matrices could not be rejected (Box’s $M = 8.66$, $F = 0.94$, $p = .49$). Once again, homogeneity of variance was established for both dependent variables, as the null hypothesis of equal variance for both dependent variables ($F(3, 156) = 1.28, p = .28$ for posttest and $F(3,156) = 1.06, p = .37$ for practicum) was not rejected at a 5% level.

Because of the non-significance of Box’s $M$, thus indicating homogeneity of covariance matrices, Wilk’s Lambda was again the appropriate test to examine the multivariate effect of type of instruction and gender on the posttest and practicum scores. Using an alpha level of .05, this test was significant for type of instruction, Wilk’s $\Lambda = 0.93$, $F(2, 155) = 5.98, p = .003$, partial $\eta^2 = .07$. The significant $F$-value indicated that the scores of the posttest and practicum were significantly dependent on which type of instruction the students received. The equivalent test for gender, however, was not significant, with Wilk’s $\Lambda = 0.99$, $F(2, 155) = 0.83, p = .44$, multivariate $\eta^2 = .01$. The partial $\eta^2$ for type of instruction indicated that approximately 7% of the variance in the dependent variables was associated with type of instruction, while 1% of the variation in the posttest and practicum scores combined was associated with gender (Horn, 2013).

In order to determine how the dependent variables varied for the independent variables, the univariate ANOVAs of the variables were examined next. According to the ANOVA table, the type of dissection instruction had a statistically significant effect on the posttest score ($F (1, 156) = 11.55; p = .001; \text{partial } \eta^2 = .07$), but not on the practicum score ($F(1, 156) = 0.02; p = .90; \text{partial } \eta^2 < .001$). Also according to the ANOVA table, the student’s gender did not have a significant effect on either the posttest
score ($F(1, 156) = 0.64; p = .43; \text{partial } \eta^2 = .004$), or the practicum score ($F(1, 156) = 0.65; p = .42; \text{partial } \eta^2 = .004$). As was previously explained, it is important to note that, in order to account for multiple ANOVAs being run, an alpha correction, such as a Bonferroni correction, should be made. With this correction a statistical significance of $p < .025$ would be considered significant; neither variable met this criterion for the practicum score, but type of instruction again was significantly related to the posttest score (Laerd Statistics, 2013). In conclusion, the MANOVA test indicated that the type of dissection instruction received by students influenced the combined grades of the posttest and the practicum. Gender did not influence the combined grades of the posttest and the practicum. However, when examined separately, it appeared that the posttest score was influenced by the type of instruction, but the practicum score was not, while neither score was influenced by gender.

**Independent samples t test, differences by ethnicity.** In order to begin to answer RQ2 with regard to ethnicity, an independent samples $t$ test was first run examining Caucasian and non-Caucasian students separately for the virtual dissection group only. Independent samples $t$ tests are employed to examine pairs of variables only; one test could not be used to investigate differences between all ethnicity variables at the same time. The tests needed to be conducted in pairs, hence Caucasian and non-Caucasian students, African-American and non-African-American students, etc. For the first pair of variables examined the question was, do Caucasian or non-Caucasian students benefit the most from the virtual dissection technique, based on their score improvement? Figure 22 shows the means for Caucasian and non-Caucasian students in the virtual dissection group. Mean scores improved for both race designations, but
appear to have increased more for Caucasian students (7.80) than for non-Caucasian students (5.50). Whether this difference was significant or not was determined by means of an independent samples $t$ test.

![Bar chart](image)

**Mean Score Improvement, Virtual Dissection Students Only**

*Overall and By Ethnicity: Caucasian and Non-Caucasian Students*

<table>
<thead>
<tr>
<th></th>
<th>Caucasian</th>
<th>Non-Caucasian</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>6.18</td>
<td>5.50</td>
</tr>
<tr>
<td>N=51</td>
<td>N=36</td>
<td></td>
</tr>
</tbody>
</table>

Figure 22. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case Caucasian and non-Caucasian students.

For this comparison, Levene’s test for equality of variances was again run to test the assumption of the corresponding $t$ test. The null hypothesis for Levene’s test (that the variance for Caucasian students was the same as the variance for non-Caucasian students) was not rejected for this data at a significance level of $\alpha = .05$, with $F(49) = 0.17$, $p = .68$. This meant that the $t$ value for equal variances was used in this case.

The null hypothesis for this $t$ test was that the mean score improvement for Caucasian students was the same as the mean score improvement for non-Caucasian students in the virtual dissection group. Using the $t$ test for equal variances yielded the result: $t(49) = -0.30$, $p = .76$. Hence the null hypothesis was not rejected. Although the score improvement for the virtual dissection group appeared higher for Caucasians than
non-Caucasians, the \( t \) test results showed that the difference was not significant at a 5\% level. The boxplot in Figure 23 shows the difference in the mean score improvement by ethnicity: Caucasians and non-Caucasians for the virtual dissection group only.

![Boxplot of score improvement variable for Caucasian and non-Caucasian students in the virtual dissection group only.](image)

Figure 23. Boxplot of score improvement variable for Caucasian and non-Caucasian students in the virtual dissection group only.

In order to continue answering RQ2 with regard to ethnicity, an independent samples \( t \) test was next run examining African-American and non-African-American students separately for the virtual dissection group only. In other words, do African-American or non-African-American students benefit the most from the virtual dissection technique, based on their score improvement? Figure 24 shows the means for African-American and non-African-American students in the virtual dissection group. The mean score improvement was slightly higher for African-American students (7.36), than non-African-American students (5.73).
Mean Score Improvement, Virtual Dissection
Students Only
Overall and By Ethnicity: African-American and Non-African-American Students

<table>
<thead>
<tr>
<th></th>
<th>All students</th>
<th>African-American</th>
<th>Non-African-American</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score Improvement</td>
<td>6.18</td>
<td>7.36</td>
<td>5.73</td>
</tr>
<tr>
<td>N</td>
<td>51</td>
<td>14</td>
<td>37</td>
</tr>
</tbody>
</table>

Figure 24. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case African-American and non-African-American students.

Note that scores improved for both race designations, and the increase was slightly higher for African-American students. Whether the slight difference is significant or not was determined by means of an independent samples $t$ test. It was again necessary to test whether the variance for the score improvement for African-American students was equal to the variance for the score improvement variable for non-African-American students for virtual dissection students only. The null hypothesis for Levene’s test (that the variance for African-American students was the same as the variance for non-African-American students) was not rejected at a significance level of $\alpha = .05$, with $F(49) = 3.26, p = .08$.

The null hypothesis for this $t$ test was that, for students in the virtual dissection group, the mean score improvement for African-American students was the same as the mean score improvement for non-African-American students. Using the $t$ test for equal variances yielded the result: $t(49) = -0.21, p = .83$. Hence the null hypothesis that the
mean score improvement was the same for African-American and non-African-American students in the virtual dissection group was not rejected. This is consistent with Figure 24, which showed the mean score improvement for African-Americans was only slightly higher than the mean score improvement for non-African-American students. The boxplot in Figure 25 shows the difference in the mean score improvement by ethnicity: African-Americans and non-African-Americans for the virtual dissection group only.

Figure 25. Boxplot of score improvement variable for both African-American and non-African-American students in the virtual dissection group only.

Finally, to answer RQ2 with regard to ethnicity, an independent samples $t$ test was run examining Hispanic and non-Hispanic students separately for the virtual dissection group only. In other words, do Hispanic or non-Hispanic students benefit the most from the virtual dissection technique, based on their score improvement? Figure 26 shows the
mean scores for Hispanic (7.85) and non-Hispanic (5.61) students in the virtual dissection group. The mean score improvement was higher for Hispanic students.

Figure 26. Mean score improvement for virtual dissection students only, broken down by ethnicity, in this case Hispanic and non-Hispanic students.

Note that scores improved for both ethnic designations, but increased more for Hispanic students than for non-Hispanic students. Whether this difference was significant was determined by means of an independent samples $t$ test. Levene's test was used to test whether the variance for the score improvement for Hispanic students was equal to the variance for the score improvement variable for non-Hispanic students in the virtual dissection group only. The null hypothesis for Levene's test (that the variance for Hispanic students was the same as the variance for non-Hispanic students) was not rejected at a significance level of $\alpha = .05$, with $F(49) = 1.81, p = .18$.

The null hypothesis for this $t$ test was that the mean score improvement for Hispanic students was the same as the mean score improvement for non-Hispanic
students in the virtual dissection group. The \( t \) test for equal variances yielded the result: \( t(49) = -0.28, p = .78 \). Hence the null hypothesis that the mean score improvement was the same for Hispanic and non-Hispanic students in the virtual dissection group was not rejected. Although the score improvement for the virtual dissection group appeared higher for Hispanics than non-Hispanics, the \( t \) test results indicated that the difference was not significant at a 5% level. The boxplot in Figure 27 shows the difference in the mean score improvement by ethnicity: Hispanics and non-Hispanics for the virtual dissection group only.

![Boxplot of score improvement variable for both Hispanic and non-Hispanic students in the virtual dissection group only.](image)

**Figure 27.** Boxplot of score improvement variable for both Hispanic and non-Hispanic students in the virtual dissection group only.

Overall, the \( t \) test results illustrated that, for the virtual dissection students, neither gender nor ethnicity played a role in influencing score improvement. According to the test statistics, there was not a significant difference between pretest and posttest scores
between males and females, or between the different ethnicities. Although the mean score did increase overall between the pretest and posttest for all students, this increase appeared to be unrelated to either gender or ethnicity.

**MANOVA analysis, differences by ethnicity.** For this question, factorial MANOVA allowed for the examination of the interaction of other variables that might influence science scores other than the type of dissection instruction, ethnicity in this instance. For RQ2, two different factorial MANOVA models with more than one independent variable each were employed to answer the question. Type of instruction was included as one of the independent variables in this model, while ethnicity was included as the other independent variable. In this model, the independent variable ethnicity yielded four options: Caucasian, African-American, Hispanic, and Other.

The Box’s M test of homogeneity of covariance matrices found that the null hypothesis of equal covariance matrices could not be rejected (Box’s $M = 19.46, F = 0.88, p = .62$). Once again, homogeneity of variance was established for both dependent variables, as the null hypothesis of equal variance for both dependent variables ($F(7, 156) = 0.95, p = .47$ for posttest and $F(7, 156) = 0.72, p = .66$ for practicum) was not rejected at a 5% level.

Because of the non-significance of Box’s $M$ indicating homogeneity of covariance matrices, Wilk’s Lambda was again the appropriate test to use to examine the multivariate effect of type of instruction and ethnicity on the posttest and practicum scores. Using an alpha level of .05, this test was significant for type of instruction, Wilk’s $\Lambda = 0.92, F(2, 155) = 6.47, p = .002$, partial $\eta^2 = .08$. The significant $F$ value indicated that the scores of the posttest and practicum were significantly dependent on
which type of instruction the students received. The equivalent test for ethnicity, however, was not significant, Wilk’s $\Lambda = 0.98, F(6, 155) = 0.64, p = .70$, partial $\eta^2 = .01$. The partial $\eta^2$ for type of instruction indicated that approximately 8% of the variance in the dependent variables was associated with type of instruction, while 1% of the variation in the posttest and practicum score was association with ethnicity.

In order to determine how the dependent variables differed for the independent variables, the univariate ANOVAs of the variables were examined next. According to the ANOVA table, the type of dissection instruction had a statistically significant effect on the posttest score ($F(1, 156) = 13.01; p < .001; \text{partial } \eta^2 = .08$), but not on the practicum score ($F(1, 156) = 0.51; p = .48; \text{partial } \eta^2 = .003$). Also according to the ANOVA table, the student’s ethnicity did not have a significant effect on either the posttest score ($F(3, 156) = 1.05; p = .37; \text{partial } \eta^2 = .02$), or the practicum score ($F(3, 156) = 0.09; p = .97; \text{partial } \eta^2 = .002$). With the Bonferroni correction, a statistical significance of $p < .025$ would be considered significant; neither variable met this criterion for the practicum score, but type of instruction again was significantly related to the posttest score (Laerd Statistics, 2013). In conclusion, therefore, the MANOVA test indicated that the type of dissection instruction received by students influenced the combined grades of the posttest and the practicum. Ethnicity, like gender, did not influence the combined grades of the posttest and the practicum. However, when examined separately, it appeared that the posttest score was influenced by the type of instruction, but the practicum score was not, while neither score was influenced by ethnicity.
Research Question 3

The third research question examined in the study was: What preference do high school biology students have as a means for acquiring the knowledge of human systems: virtual dissection of organisms or the traditional dissection method? All students in the classes involved in the intervention were pre-tested and then separated into two groups. Some did the virtual activity while others did the hands-on, traditional dissection. Then they were post-tested. After the posttest and practicum, the groups were switched and each student performed the other type of activity (virtual or traditional dissection). Afterward, students were asked to complete a survey about the interventions. In the survey, they were asked which type of dissection experience they preferred. The results of the survey are shown in Figure 28. The survey results indicated nearly 66.5% of the students preferred the traditional dissection method.

![Type of Instruction Preferred by Students](image)

Figure 28. Pie chart showing which type of dissection instruction students preferred.
In order to gain a greater understanding into whether students were merely choosing the method they used originally, a cross tabulation examining the frequency of students in each instruction group who preferred each type of dissection instruction. Figure 29 shows this breakdown of preference by type of instruction originally received. The results indicated that students who originally experienced the traditional rat dissection method greatly preferred this method over the virtual method 69.5% to 30.5%. Of those students who originally experienced virtual rat dissection, 60% of the students still preferred the traditional method over the virtual method which had a preference percentage of 40%.

Figure 29. Type of dissection preferred by type of instruction originally received.
Summary

This study was based on an examination of the pretest, posttest, and practicum scores of high school students who participated in either a traditional or virtual rat dissection lesson. The pretest was administered to all students before the lesson regardless of whether they were assigned to the traditional or the virtual dissection lesson. At the completion of the lesson, all students were then administered a posttest and a practicum test. In addition, once the scores were recorded, the students also participated in the other type of dissection instruction for comparison. At that point, the students were asked to complete a survey indicating which method they preferred. Descriptive statistics were shown in this chapter, including an overview of the demographics of the students and comparisons of the means of the various groups.

The first research question, which examined the effect of dissection method on student achievement, was investigated using both an independent samples t test and a one-way MANOVA analysis. According to the t test analysis, improvement of the mean test score was not the same for both dissection groups; the mean score improved significantly more for the group that participated in traditional dissection. The MANOVA analysis indicated that type of instruction did influence posttest and practicum scores directly; scores were significantly higher in the traditional dissection group.

The second research question was an examination of the effect on student learning of virtual dissection instruction on students of different gender and ethnicity. This question was also analyzed using both an independent samples t test and a factorial
MANOVA model with two independent variables. Overall, the $t$ test results illustrated that, for the virtual dissection students, neither gender nor ethnicity played a role in influencing score improvement. The MANOVA test indicated that the type of dissection instruction received by students influenced the combined grades of the posttest and the practicum, but the results were not influenced by gender or ethnicity.

Research question 3 was intended to examine which type of instruction was preferred by students, based on the survey results. Over $\frac{2}{3}$ (66.5\%) of the students surveyed preferred the traditional dissection method. In conclusion, not only did the traditional dissection method appear to yield greater gains from pretest to posttest scores than the virtual method (as demonstrated in RQ1 and RQ2), but students also tended to prefer this type of dissection method. In Chapter 5, the implications of this study are discussed and recommendations for additional research are suggested.
CHAPTER 5: DISCUSSION

The purpose of this chapter is to discuss the impacts of this study in the context of the effectiveness of virtual dissection versus the traditional hands-on method in improving student achievement. It begins with a summary of the findings, which is followed by conclusions drawn from the results. Then the limitations of the study are presented, as well as potential directions for future research on the topic. Finally, assessment of the practical implications of the study concludes this work.

Summary of Findings

The purpose of this mixed-methods study was to examine the effectiveness of instruction through virtual dissection, when compared to traditional hands-on rat dissection, on student achievement. Three research questions guided this investigation:

1. What effect on student achievement do virtual animal dissection and traditional dissection methods have on a general animal anatomy assessment (consisting of recall questions), and a lab practicum exercise (consisting of organ identification and function)?

2. What effect does virtual dissection (as opposed to the traditional method) have on student achievement, focusing on gender and the primary ethnicities (Caucasian, African-American and Hispanic)?

3. What preference do high school biology students have as a means for acquiring the knowledge of human systems: virtual dissection of organisms or the traditional dissection method?

For the first research question, the differences in student achievement between a traditional hands-on dissection method and a virtual dissection method were examined in
detail. This research question was investigated using both an independent samples $t$ test and one-way MANOVA analysis. ANOVA, as well as MANOVA, results were reported in Chapter 4. A quick examination of the mean posttest and practicum scores showed that the scores appeared to increase more for students taught using traditional rat dissection than for students who experienced the virtual dissection. Posttest scores did however increase over pretest scores in both groups of students.

An independent samples $t$ test was conducted: this tested the hypothesis that the mean score improvement between the pretest and posttest was the same for both groups. The results led to a rejection of the null hypothesis, with $p=.001$. The mean score improvement was not the same for both dissection groups; it was significantly greater in the traditional dissection group. Using a newly created dependent variable based on a combination of the posttest and practicum scores, a MANOVA analysis was then used to examine the influence of the type of dissection instruction on the two dependent variables, posttest and practicum score. The type of instruction (traditional or virtual) had a statistically significant influence on this combination variable, with $p=.005$, meaning that the combined scores were significantly higher for the traditional instruction group. Additionally, the MANOVA results showed that 6% of the variance in the dependent variable was associated with the type of instruction. An examination of the univariate ANOVAs of the variables showed that only the posttest score was actually influenced by type of instruction with $p=.001$; practicum scores were not influenced by instruction type with $p=.69$.

For research question 2, the effect of virtual dissection (as opposed to the traditional method) on student achievement differences was examined with the focus on
gender and the primary ethnicities (Caucasian, African-American, and Hispanic). As with Research Question 1, this question was analyzed using both an independent samples $t$ test and a MANOVA model. In this case, a factorial MANOVA with two independent variables was used.

Here only results from the virtual dissection group were analyzed in the $t$ test. The researcher tried to determine whether either gender or ethnicity, or both, influenced posttest or practicum scores of students who experienced virtual dissection. Did students of one gender or one particular ethnicity respond more positively to the virtual dissection environment? Overall, the $t$ test results illustrated that within the virtual dissection group, neither gender nor ethnicity played a role in influencing score improvement. Score improvement (from pretest to posttest) of male students in the virtual dissection group was not significantly different from score improvement of female students, with $p=.48$. Similarly, score improvement of Caucasian students were not significantly different from score improvement of non-Caucasian students, with $p=.76$. The same held true for African-American students when compared to non-African-American students ($p=.83$) and Hispanic students when compared to non-Hispanic students ($p=.78$). According to the reported test statistics, there was not a significant difference in score improvement from pretest to posttest between males and females, or between the different ethnicities.

Analyzing the factorial MANOVA model with gender included as an independent variable illustrated that as with Research Question 1, the type of instruction did influence the joint combination of posttest and practicum scores, with $p=.003$. However, the gender variable was not statistically significant, with $p=.44$. The factorial MANOVA model with ethnicity as an independent variable yielded similar results, with type of
instruction significantly related to the dependent variable with $p=.002$ and ethnicity not statistically significant with $p=.70$. Including gender and ethnicity in the models did not influence the combination dependent variable; all gender and ethnicity results were not significant at a 5% level. A breakdown of the univariate ANOVAs of the variables yielded similar results to those reported for research question 1: posttest scores were influenced by type of instruction while practicum scores were not. This held true for both the gender and the ethnicity factorial MANOVA models.

Finally, research question 3 addressed whether the high school biology students in the sample had a preference for the virtual dissection of organisms or the traditional dissection method. All students in the participating classes were pre-tested and then separated into two groups, after which some students did the virtual activity while others did the hands-on traditional dissection. After this dissection activity, all students were post-tested. After the posttest and practicum, the groups were switched and each student performed the other type of dissection activity. Students were then asked to complete a survey indicating which type of dissection experience they preferred.

According to the survey results, 66.5% of students preferred the traditional method while 33.1% preferred the virtual method. By examining these responses, and taking into consideration which type of activity each student tried first (and was tested on), it was determined that 69.5% of students in the traditional dissection group preferred the traditional dissection method. Similarly, 60% of students in the virtual dissection group preferred the traditional dissection method.
Conclusions

The results from the analysis of RQ1 indicated that improvement between pretest and posttest scores was higher for students in the traditional rat dissection group as compared to the virtual dissection group. The results of this analysis also indicated that type of dissection instruction did influence posttest and practicum scores directly. While the one-way MANOVA test indicated that scores were significantly higher for the traditional dissection group, the univariate ANOVA breakdown showed that only the posttest score was actually influenced by type of dissection instruction; practicum scores were not influenced by instruction type. Overall, the analysis supported the conclusion that students performed better when taught using traditional rat dissection methods rather than a virtual rat dissection method.

The second research question addressed whether the results of RQ1 varied by gender or by ethnicity. According to the statistical analysis done on the virtual dissection group only, there was not a significant difference between pretest and posttest scores of male and female students, or between the different ethnicities. So although the overall mean score for all students increased between the pretest and posttest, this increase was unrelated to either gender or ethnicity. Factorial MANOVAs were then performed with gender and ethnicity, respectively, as independent variables. The results of these MANOVA tests were consistent with the results of RQ1; type of instruction did influence the dependent variable, combined posttest and practicum scores. When examined separately in both the gender and ethnicity models, only the posttest score was influenced by the type of instruction (not the practicum score) and neither score was influenced by gender or ethnicity. In conclusion, the results of RQ 1 were consistent across genders
and ethnicities, where students taught using the traditional dissection method consistently scored higher than students taught using the virtual dissection method; however gender and ethnicity were not significant variables in these models.

In order to answer RQ3, students were asked to state their preference for instruction type in a survey. Over 2/3 (66.5%) of all the students surveyed preferred the traditional dissection method, including a number of students who were originally taught by the virtual dissection method. In summary, not only did the traditional dissection method appear to yield greater gains from pretest to posttest scores than the virtual method (as demonstrated in RQ1 and RQ2), but students also tended to prefer this type of dissection method.

The results discussed here are generally consistent with other studies that have been conducted in this field. Michael-Clark (2003) found that students in a hands-on frog dissection group scored significantly higher on assessments than a virtual activity group. This author also found those results to be consistent across genders and ethnicities. Cross & Cross (2004) also found that a hands-on frog dissection group outperformed a virtual frog dissection group at a statistically significant level. The independent samples t tests and the MANOVA analyses carried out for RQ1 and for RQ2 of this study support the conclusions drawn by Michael-Clark (2003) as well as Cross & Cross (2004).

Another finding of this study was that when the univariate ANOVA results from the MANOVA procedure used in RQ1 and RQ2 was analyzed in detail, it was seen that the practicum results were not significantly different between the virtual and traditional dissection groups. This could indicate that the score advantage of the traditional group over the virtual group may not be as great as it seemed from the MANCOVA.
Montgomery (2008) and Kopec (2002) found no significant differences in scores between a traditional and a virtual dissection group. Therefore more research is needed to verify this specific result.

The current study indicated that the traditional dissection method of instruction was more effective in teaching students and was also preferred by students. It must be noted that this is just one study and as such, conclusions and generalizations should not be drawn from the results of this study alone. The study would need to be expanded and replicated, preferably many times, before acting on the conclusions drawn here.

Additional limitations of this study also existed and are presented in the next section.

**Limitations**

Although limitations to this study existed, a number of strengths should also be recognized. Strengths of this work included the relatively large sample size. The sample was also quite diverse; it was fairly evenly divided between male and female students, and also between Caucasian, African-American, and Hispanic students. A pretest posttest design was used in order to compare learning outcomes from the dissection lesson; administering a pretest helped to compensate for any initial differences in knowledge between students. Thus learning was measured, rather than innate knowledge of the subject. In addition, many previous studies have focused on virtual frog dissections (Apkan & Andre, 2000; Kinzie, Strauss, & Foss, 1993; Sweitzer, 1996), however this study was the first of its kind that analyzed rat dissection at the high school level. Rat dissection specimens were chosen for use here, as rats are believed to provide the students a closer analysis of an anatomy to that of the human system. The
methodology used in the analysis was sound and the findings were consistent with previous studies.

Despite the strengths described above, this study also had limitations that must be acknowledged. Although there was a relatively large sample, the groups within the sample were unequal in size. The traditional dissection group was over 1.5 times larger than the virtual dissection group. Statistical tests are more reliable when group sizes are similar, so the results would have been stronger if the samples had been more equal in size. Additionally, the overall sample size was large, the number in each gender and race category was quite small. Because of the small values of N (especially in the virtual dissection group that was analyzed in the t tests related to RQ2), any results from the gender and ethnicity models should not be generalized to larger populations and samples.

It is also recognized that there could have been a school and/or teacher effect occurring in this study. The sample was limited to a single, mid-size high school in a district located in Central-South Florida. The biology teachers at this high school participated voluntarily. The teachers may have conducted the study at slightly different times, thus possibly leading to sharing of information between students. The participants (students) were not in a controlled environment, meaning that the results of this study cannot be interpreted as indicating a correlation between time and test results. Information sharing between students could also have influenced students' preferences on the survey, as some students may have been influenced by others' opinions.

Similarly, there was no one single instructional approach used by teachers. Animal dissection presentations can be implemented in a variety of ways and the teachers were encouraged to deliver the information deemed most appropriate for their students.
Consequently, this meant that there was not one approach to the presentation of the material; thus this study could not reveal connections between student achievement and mechanisms of implementation. Differences among teachers and teaching methods could have influenced the results.

Since this study was isolated to one school, generalizability was also limited. The students in this study were enrolled in a suburban public high school. The results may not be used to generalize to those obtained from a public inner city high school or private high school setting.

Lastly, the only information given on the students' academic backgrounds was whether they previously passed physical science or not. Students who did not pass this course were not included in the sample, according to the criteria set forth in Chapter 1. For this particular study, given the information available, it was impossible to break down the students to see which type of student benefitted most from traditional instruction and which type may have benefitted from virtual instruction. No information was available on each student's GPA. It is possible that lower GPA students benefitted most from the virtual dissection method, or vice versa, but that conclusion cannot be made based on the results of this work.

**Recommendations for Future Research**

The following recommendations offer suggestions for future research that could improve the understanding of the relationship between learning and virtual dissection as a teaching tool or method. Some of the recommendations address the need for enhancing and replicating the current study while others suggest alternative methodologies that could be employed to strengthen the results of this research. Suggestions also emphasize
additional types of data and programs that could be gathered and implemented in order to gain an even greater insight into these diverse interrelationships. Some ideas for future research are as follows:

- Replicate the study. The study must be repeated in order to draw any definitive conclusions from the results. It should also be enlarged to include more students, as well as students at different schools in different settings in order to improve generalizability. A larger sample size will increase the power of the study.

- Provide a script, through the use of a comprehensive dissection manual. This would provide consistency in the delivery of instruction. Since, there was no one single instructional approach used by teachers, this study could not reveal connections between student achievement and mechanisms of implementation. The differences among teachers and teaching methods could have influenced the results.

- Expand the study to also include school information on the students who participate. For example, their GPAs would be very useful in determining which type of students benefit most from virtual programs. A more in-depth study could then be conducted by breaking students down based on individual GPA and looking for anomalies.

- Other methodologies could be used for the study. One example would be a paired t test that examined the pretest and posttest scores for students in the virtual group, as well as for students in the traditional group. This test and other similar tests could be carried out on the current data, as well as on
additional data. Another methodology that could be used in this type of analysis is the ANCOVA model, which examines the posttest scores of both dissection groups and includes the pretest score as the covariate. A similar MANCOVA model could include the two dependent variables, with type of instruction as an independent variable (as well as possibly gender or ethnicity) and pretest scores as the covariate.

- As the practicum stands, with only three questions, it does not lend much additional depth to the research. A more involved practicum test should be developed; this would greatly help the ANOVA analysis. Conversely, the practicum could be dropped and the research based solely on the posttest, or the practicum could be integrated into the posttest to create one comprehensive test. Another possibility would be to administer a pre-practicum test, in order to gain a greater understanding of which type of dissection instruction leads to higher gains on practicum scores.

- Multiple regression analysis could be used with the posttest (or a scale variable combining both the posttest and the practicum score) as the dependent variable. The major predictor variable of interest would be a dummy variable representing the type of classroom instruction. The pretest score would be a control variable. Gender and ethnicity would also be used as control variables. Demographic variables such as family income and disability would add more depth to the model. School data such as GPA and science GPA could also be important control variables. The coefficient on the classroom instruction variable would then be examined for sign and
significance. This test could be run on different classes and in different schools, again looking for consistency of results among the different settings.

- Investigate other types of virtual dissection programs. Likely, more than one rat dissection program is available for students to use. The results with the program used in the current research were not encouraging at first glance, however other programs may yield more encouraging results.

- The specific problems with the virtual program could also be examined in more detail. On the survey, students indicated that they preferred the traditional dissection method to the virtual method. The survey did not delve into why students preferred one method over another. Additional research should include the development of a more detailed exit survey for administering to students. Students could then indicate more specific reasons for their preferences; these results could then be used to develop virtual dissection programs that may yield more positive results.

**Implications for Practice**

Controversy currently exists concerning the use of traditional dissection methods in the classroom. Opponents of traditional dissection methods often focus on the ethical issues concerning the use of animals for dissection (Russell, 1996; Balcombe, 1996, 1997, 1998, 2000, 2001). Other researchers focus on the benefits of traditional dissection as a necessary part of the educational experience (Morrison, 1992; Russo, 1997; Berstein, 2000; Valli, 2001). Proponents of virtual dissection cite benefits of this teaching method including the advantage of being able to repeat the lesson (Hepner, 1996) and the non-destructive nature of virtual dissection lessons (Rose, 1995). However, decisions to
replace traditional, hands-on dissection with virtual dissection in the classroom must be made carefully. Although the current study illustrated that, in the sample from a Florida high school, students taught using traditional rat dissection methods outperformed students taught using a virtual method, it is imperative not to reach too far in generalizing or acting upon these results. The implications of this work for practice in this field must take into consideration the limited generalizability of the study and the need for more research in this area.

It is important to remember that the current research compared mean scores. Without pairing, there was no way of knowing whether, for a subset of students, the augmented reality method might not have worked better. This point should be emphasized because both methods resulted in learning. It is possible that some types of students learned more using the augmented reality method than they would have if in a live classroom. For example, both regular and honors biology students were included in the sample. The data was not broken down to show which students were in the honors classes and which were not. It is possible that students in the honors classes performed either better or worse in the virtual setting than other students. However, from the current study, any determination such as this is impossible to make.

This one study indicated that the traditional dissection method of instruction was more effective in teaching students and was also preferred by students. As mentioned before, generalizability is limited, and the study would need to be expanded and replicated before acting on its conclusions. Too often education studies like this one are taken as definitive, and educators eliminate a form of teaching in favor of the one with the higher mean score. Mean scores are useful for administrators and teachers to evaluate
their performance, but relatively useless to individual students, not one of whom is likely to have the mean score. The mean is an average and does not apply to every student. It would be a mistake to conclude here that live classrooms are always better. Since both methods resulted in an increase of mean scores from the pretest to the posttest, it is imperative as a follow-up to perform a paired $t$ test and to create a CHAID (chi-squared automatic interaction detector) decision tree to gain a better understanding of how individual students learn. More work needs to be done, even with this data set, before actions such as eliminating virtual dissection as a teaching method for all students are undertaken.

Furthermore, individuals that influence educational policy need to continue to consider what is best for the student (not solely cost effectiveness) in helping drive student achievement. Far too often states’ policy on education is predicated on the rise and fall of their economy. If the economy of a state takes a downturn, cuts to a state’s budget (including the educational portion of the budget) are implemented. This forces school districts to make provisions to continue to provide adequate education with fewer resources. Based on this research and others relating to classroom activities, student achievement, and educational funding, it is the belief of this author that policy makers should remain proactive in funding their state’s public education budget in order for districts to deliver differentiated instruction and activities, thus maximizing student achievement.

The author’s belief and this study align itself with the role of The Carnegie Project on the Education Doctorate (CPED) consortium that Lynn University is a member of. The intent of CPED is to collaboratively redesign the Ed. D. and to make it a
stronger and more relevant degree for the advanced preparation of school practitioners and clinical faculty, academic leaders and professional staff for the nation’s schools and colleges and the learning organizations that support them. (Carnegie Project for the Education Doctorate, 2014) The consortium has created an education doctorate definition and working principles that are used to guide the development of quality professional practice preparation while supporting the diverse needs of local context within colleges’ of education. The six (6) defining principles help to set research and development agendas to test, refine, and validate principles for the professional doctorate in education. This study supported the first principle:

1. *Is framed around questions of equity, ethics, and social justice to bring about solutions to complex problems of practice.* (Carnegie Project for the Education Doctorate, 2014)

Focusing on differentiated activities in relation to student achievement, this study brought to the forefront equity within the classroom, targeting the on the optimal way for a student to learn content.

The findings of this study primarily indicate that more research still needs to be carried out in this area. The results provided some evidence that traditional hands-on rat dissection methods are more effective than virtual methods in imparting knowledge to students; again however, this result should not be generalized or acted upon without further research. Much more research is needed before any conclusions about replacing traditional dissection lessons with virtual dissection lessons can be drawn. Such factors as different virtual dissection programs, different classes, and different school settings
could easily yield different results. Similar studies using different virtual programs in different schools in different geographical areas should be carried out before acting upon the results of the current study. Although in this study, virtual dissection instruction appeared less effective than traditional hands-on instruction, other studies may yield very different results.
REFERENCES


http://oak.ucc.nau.edu/rh232/courses/EPS625/Handouts/Interpreting%20the%20One-Way%20MANOVA.pdf


Maloney, R. (2002). Virtual fetal pig dissection as an agent of knowledge acquisition and attitudinal change in female high school biology students.


Michel-Clark, I. (2003). A comparison of the effects of animal dissection and a computer simulation dissection program on students' knowledge of frog anatomy and attitudes toward dissection (Doctoral dissertation). Retrieve from ProQuest Dissertations and Theses Database. (UMI No. 3090895)


Office of the Press Secretary. (2010). *President Obama expands “educate to innovate” campaign for excellence in science, technology, engineering and mathematics education*. The White House.


Retrieved October 5, 2012 from MasterFILE Premier database.


St. Lucie Public Schools. (2012). Full-time equivalent student status [Data file].


The Royal Society (n.d.) Retrieved December 18, 2012 from

http://royalsociety.org/about-us/


STUDENT NUMBER: ____________  GENDER (CIRCLE ONE): MALE
FEMALE

RACE (CIRCLE ONE): CAUCASIAN  AFRICAN-AMERICAN  HISPANIC
OTHER

Did you pass your physical science class last year (circle one)?  YES  NO

Rat Dissection Pretest

1. What are the 4 levels of organization?
   a. Tissue, cell, organ, organ system.
   b. Cell, organ, tissue, organ system
   c. Cell, tissue, organ, organ system
   d. Organ, organ system, cell, tissue

2. What are the 4 major organ systems?
   a. __________________________
   b. __________________________
   c. __________________________
   d. __________________________

3. In order, list the organs of the digestive system?
   a. __________________________
   ______

4. What is the function of the Esophagus?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

5. What is the function of the mouth?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients
6. What is the function of the Stomach?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

7. What is the function of the intestine?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

8. In order, write the steps oxygen moves through the respiratory system.
   a. __________________________

9. What is the function of the nose?
   a. Inhale and warms the oxygen
   b. Provides a pathway between nose and trachea
   c. Provides a pathway between pharynx and lungs
   d. Exchanges O₂ and CO₂ from blood

10. What is the function of the Trachea?
    a. Inhale and warms oxygen
    b. Provides a pathway between nose and trachea
    c. Provides a pathway between pharynx and lungs
    d. Exchanges O₂ and CO₂ from blood

11. What is the function of the lungs?
    a. Inhale and warms oxygen
    b. Provides a pathway between nose and trachea
    c. Provides a pathway between pharynx and lungs
    d. Exchanges O₂ and CO₂ from blood

12. What is the function of the Pharynx?
    a. Inhale and warms oxygen
    b. Provides a pathway between nose and trachea
    c. Provides a pathway between pharynx and lungs
    d. Exchanges O₂ and CO₂ from blood
13. Which is NOT a function of the heart?
   a. Pumps blood into the lungs
   b. Keeps us alive
   c. Maintain homeostasis
   d. Filters waste out of the blood.

14. Which chambers of the heart carry oxygenated blood?

15. Which chambers of the heart carry deoxygenated blood?

16. In order, write the steps waste moves through the excretory system.

17. What is the function of the kidney?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body

18. What is the function of the urinary bladder?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.

19. What is the function of the urethra?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.

20. What is the function of the ureter?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.
Rat Dissection Posttest

1. What are the 4 levels of organization?
   a. Tissue, cell, organ, organ system.
   b. Cell, organ, tissue, organ system
   c. Cell, tissue, organ, organ system
   d. Organ, organ system, cell, tissue

2. What are the 4 major organ systems?
   a. ______________________
   b. ______________________
   c. ______________________
   d. ______________________

3. In order, State the organs of the digestive system?
   a. ______________________________________

4. What is the function of the mouth?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

5. What is the function of the Stomach?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

6. What is the function of the intestine?
   a. Mechanically breaks down of food
   b. Provides a pathway between mouth and stomach
   c. Mechanically and chemically breaks down food
   d. Chemical digestion and absorption of nutrients

7. In order, write the steps oxygen moves through the respiratory system.
   a. ______________________________________
8. What is the function of the nose?
   a. Inhale and warms the oxygen
   b. Provides a pathway between nose and trachea
   c. Provides a pathway between pharynx and lungs
   d. Exchanges O₂ and CO₂ from blood

9. What is the function of the Trachea?
   a. Inhale and warms oxygen
   b. Provides a pathway between nose and trachea
   c. Provides a pathway between pharynx and lungs
   d. Exchanges O₂ and CO₂ from blood

10. What is the function of the lungs?
    a. Inhale and warms oxygen
    b. Provides a pathway between nose and trachea
    c. Provides a pathway between pharynx and lungs
    d. Exchanges O₂ and CO₂ from blood

11. What is the function of the Pharynx?
    a. Inhale and warms oxygen
    b. Provides a pathway between nose and trachea
    c. Provides a pathway between pharynx and lungs
    d. Exchanges O₂ and CO₂ from blood

12. 18. Label the heart diagram.

19. Which is NOT a function of the heart?
    a. pumps blood into the lungs
    b. keeps us alive
    c. maintain homeostasis
    d. filters waste out of the blood.
20. Which chambers of the heart carry oxygenated blood?

21. Which chambers of the heart carry deoxygenated blood?

22. What is the function of the kidney?
   e. To transport urine to the urinary bladder
   f. To filter excess waste from the blood
   g. To transport urine out of the body
   h. To hold excess urine until it is released from the body

23. What is the function of the urinary bladder?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.

24. What is the function of the urethra?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.

25. What is the function of the ureter?
   a. To transport urine to the urinary bladder
   b. To filter excess waste from the blood
   c. To transport urine out of the body
   d. To hold excess urine until it is released from the body.
APPENDIX C

STUDENT NUMBER: ____________  GENDER (CIRCLE ONE):  MALE  FEMALE

RACE (CIRCLE ONE):  CAUCASIAN  AFRICAN-AMERICAN  HISPANIC  OTHER

Did you pass your physical science class last year (circle one)?  YES  NO

Practicum for Rat Dissection

There will be three questions for every organ that has to be identified. There will be the following organs: heart, lung, & stomach.

1. What is the name of the organ?

2. What main organ system does it belong too?

3. What is its main function within that system?
APPENDIX D

STUDENT NUMBER: _______________  GENDER (CIRCLE ONE): MALE  FEMALE

RACE (CIRCLE ONE): CAUCASIAN  AFRICAN-AMERICAN  HISPANIC  OTHER

Did you pass your physical science class last year (circle one)?  YES  NO

Rat Dissection Survey Questions

1. Which activity did you complete first?
   Circle one: traditional hands-on dissection / virtual dissection

2. Answer the sections about completing the traditional hands-on rat dissection.
   Describe what you did during this dissection activity:

   Like:

   Dislike:

3. Answer the sections about completing the virtual rat dissection.
   Describe what you did during this dissection activity:

   Like:

   Dislike:

4. Which activity did you prefer in acquiring the maximum knowledge of human systems – the traditional hands-on rat dissection or the virtual rat dissection? Why?
APPENDIX E

LYNN UNIVERSITY
CONSENT/ASSENT TO PARTICIPATE IN A RESEARCH STUDY

TYPE OF STUDY: Assessing High School Biology Academic Achievement by Comparing Traditional versus Virtual Dissection of Rat Specimens

INVESTIGATOR: Principal Investigator - Arthur L. Jamison, Jr.

Student, Parent/Guardian,

The purpose of this research study:
You are being asked to participate in a research study. The purpose of this study is to see if there is a difference in how well students learn rat anatomy when using a virtual rat dissection program versus the traditional hands-on method. You will also be requested complete a survey after the activity that will ask your feelings towards both activities. We want to see if there is a difference in achievement and preferred method between students who dissect a real rat and those who complete the dissection virtually.

Participants:
You are being asked to participate because you are a high school biology student. You and your parent/guardian must sign this consent/assent form in order for you to participate in the study.

Procedures:
Leading into the study, the students will choose a random number (out of a bag: 1-400) which the student will use on all documents. Then, all students will be given a pre-test to determine foundational knowledge on the anatomy of a rat specimen. The scores will be recorded and filed. Next, the teachers of record will divide each class in half based on alphabetical order (this is due to the limited amount of computers). Initially, one group will dissect a real rat while the other will study the rat using the virtual dissection program. One day before the traditional hands-on dissection, that specific group will be introduced to the safety procedures students are required to follow during the real dissection. Both dissection groups will dissect for one day. Each dissection class period will last 90 minutes. On the second day (block schedule), you will be tested on your knowledge of the rat anatomy after completing the traditional hands-on dissection or virtual dissection activity. The assessment will test your knowledge of the anatomical structures on a real rat specimen. The second set of scores will be recorded and used to determine if there were any achievement differences between the traditional hands-on group and the virtual dissection group. On the third day the two groups will switch and perform the opposite activity still utilizing the rat as the research specimen. Finally, after completing both activities, the students will be asked to fill out a survey questionnaire on their preference method for learning the content.

Participant’s Initials _______
LYNN UNIVERSITY
CONSENT/ASSENT TO PARTICIPATE IN A RESEARCH STUDY

TYPE OF STUDY: Assessing High School Biology Academic Achievement by Comparing Traditional versus Virtual Dissection of Rat Specimens

INVESTIGATOR: Principal Investigator - Arthur L. Jamison, Jr.

Risks: There are no known risks involved; only those students who sign and have a parent/guardian signature on the consent/assent form will be able to participate in the study and have their data used. Your grade will not be affected in any way by choosing to participate or refusing to participate in the study. In order to ensure confidentiality, students will choose a random number (out of a bag; 1-400) which the student will use on all documents. These values will be kept in a locked cabinet for the duration of the study. The investigators will not have access to the values.

Confidentiality: Your identity will be protected to the extent allowed by law. You will not be personally identified in any reports or publications that may result from this study. The only people who will have access to the data collected will be the Lynn University Institutional Review Board, the investigator, and the teacher of record. In order to ensure confidentiality, you will select a number. This numerical value will be used on all documents. The teacher of record will keep all relevant documents (which the numerical value will substitute for the identity of a student) locked in a cabinet for the duration of the study. The investigator will not have access to these values.

Since the numerical value ensuring a student’s confidentiality will be used on all documents, the investigator will grade the test after the teacher of record administers the tests and collects the test papers. After completion of the study, the investigator will keep the collected data for five years after which all data will be shredded by the investigator.

Participant’s Initials _______
LYNN UNIVERSITY
CONSENT/ASSENT TO PARTICIPATE IN A RESEARCH STUDY

TYPE OF STUDY: Assessing High School Biology Academic Achievement by Comparing Traditional versus Virtual Dissection of Rat Specimens

INVESTIGATOR: Principal Investigator - Arthur L. Jamison, Jr. [Redacted]

Right to Refuse or Withdraw: You may refuse to participate or withdraw from the study at anytime without penalty. Refusing to participate or declining to participate in this study will in no way harm your grade or harm your relationship with your teacher.

Questions: Any further questions you have about this study, or your participation in it, either now or any time in the future, will be answered by Mr. Arthur L. Jamison, Jr., Principal Investigator who may be reached at [Redacted] or Dr. Craig Mertler faculty advisor who may be reached at [Redacted] For any questions regarding your rights as a research subject you may call Dr. Theodore Wasserman, Chair of the Lynn University Institutional Review Board for the protection of Human Subjects at [Redacted] If any problems arise as a result of your participation in this study, please call the Principal Investigator Arthur L. Jamison, Jr. or the faculty advisor Dr. Craig Mertler immediately. Please print a copy of this consent form for your records.

Signature of Participant Date

Signature of Parent/Guardian Date

Signature of Investigator Date

Signature of Witness Date

Participant’s Initials _____
Good morning/afternoon! My name is Mr. Arthur Jamison, Jr. and I’m going to speak to you about a research project that is going to be conducted here at Fort Pierce Central High School. Your class has been selected to participate in the study. In a few weeks, your biology class will conduct a rat dissection by both the traditional hands-on method (that is where you are actually cutting a dead preserved rat) and by virtual computer simulation. The virtual simulation program that will be used is called E-Rat. Everyone in your class will complete assignments related to this unit, but only those of you who agree to participate and have your parent’s consent will participate in the actual research activities related to the unit. None of your names will be used in the research study (dissertation), so no one will know who you are. Your participation is voluntary. Your teacher will not see any of the data, except for the test scores. You can also withdraw from the study at any time.

I will hand out the consent/assent forms and will review the forms with you. Please ask me any questions you might have about the study or the consent/assent form. You are asked to sign the consent/assent form and have a parent or guardian sign the form. A parent/guardian must sign the consent assent form in order for you to take part in the study (any further questions about this study, or your participation in it, either now or any time in the future, will be answered by Mr. Arthur Jamison, Jr., Principal Investigator who may be reached at [contact information removed]. You will also be given a copy of the signed consent/assent form to keep for your records. You need to turn the signed form in to me no later than [due date]. I will be coming into your class everyday until the due date to collect the consent/assent forms. They will be kept in a sealed envelope and in a locked file cabinet. I will now review the consent/assent form with you.