7-16-2015

The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives

Ryan L. Sittler
Indiana University of Pennsylvania

THE EFFECTS OF AUDIO AND GENDER WITHIN A 3D GAMING ENVIRONMENT ON THE ACHIEVEMENT OF DIFFERENT EDUCATIONAL OBJECTIVES

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Ryan L. Sittler

Indiana University of Pennsylvania

May 2015

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Communications Media

We hereby approve the dissertation of

Ryan L. Sittler

Candidate for the degree of Doctor of Philosophy

Mary Beth Leidman, Ed.D.
Professor of Communications Media,
Chair

Jay Start, Ph.D.
Associate Professor of Communications Media

Zachary Stiegler, Ph.D.
Associate Professor of Communications Media

ACCEPTED

Randy L. Martin, Ph.D.
Dean
School of Graduate Studies and Research

Title: The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives

Author: Ryan L. Sittler

Dissertation Chair: Dr. Mary Beth Leidman

Dissertation Committee Members: Dr. Jay Start
Dr. Zachary Stiegler

This study examined the effects of audio and gender in a 3D gaming environment on the achievement of different educational objectives. A sample of undergraduate Communications Media students from a mid-sized university in Western Pennsylvania were recruited to take part in the study during Fall 2014. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music via Spotify).

The study utilized a 2 X 3 factorial post-test only design with two independent variables: audio (3 levels – none, default, and self-selected) and gender (2 levels – male and female). The dependent variable measured achievement on three criterion-referenced tests (identification, terminology, and comprehension) as well as an aggregate score for each group. The tests were based on content created by Dwyer and Lamberski (1977) as adapted by Almeida (2008).

Information processing theory, as well as research on audio in educational media, suggests that participants that listened to music while engaging with the educational game utilized in this study would underperform on achievement tests as compared to participants that received no audio treatment. No significant difference was found between subjects that received no audio treatment and subjects that listened to Mozart – though the Mozart group outperformed the Control on composite scores by 7.77% (out of 100%). Both the Control and Mozart group outperformed the self-selected group significantly. In addition, overall, females outperformed males on achievement tests. Further analysis discovered that participants self-identifying as a gamer or non-gamer can have an effect on performance. Overall, the study found that both music and gender do play a role in participants' achievement on criterion-referenced tests as they relate to instructional content.

ACKNOWLEDGMENTS

There's a part of me that still doesn't believe that I am writing this section of my dissertation. This is undoubtedly my finest accomplishment and yet… I have a sneaking suspicion that it will be months before I fully comprehend that this escapade is completed. It's been a long—and challenging—6 years. I expected to be finished with this process a long time ago and, as it turns out, that wasn't in the cards for me. As I've spoken with more and more people that completed this type of academic adventure… I've realized that my journey is not unique. So for my brothers and sisters that are undergoing the trials of dissertating right now—if you happen to stumble across this document—please have hope. You will finish. Just keep standing up when you get knocked down and keep putting one foot in front of the other until you've crossed the finish line.

I've started my acknowledgments with these words of encouragement because, more times than I care to admit, I needed to hear these very things from other people. Perhaps no one will read this section but me. And if not, well, the catharsis is nice. But nonetheless, I thought it only appropriate to share with you what those other people had shared with me. And with that… I'm going to let them now stand up, take a bow, and be recognized – along with all of the other people that have aided me along the way.

I first need to thank my wonderful wife Heatherlee J. Sittler. Her patience, support, empathy, and enthusiasm have been immeasurable. Every time I was ready to quit she was there to help me pick myself back up off the ground, get dusted off, and get me moving in the right direction again. She also listened to a whole lot of cursing and didn't chide me for it once. There's so much I could say here but words will never capture it. So, I'll just say: Heatherlee—

my statistically significant other—I love you. Thank you – and "I don't care what they say. Those numbers don't mean a thing. We've got it made" (Aukerman, 1996, track 11).

Since I'm discussing family, my mother and father have both been steadfast in their support of my academic endeavors. This time was no exception. I was the first kid in my family to go to college and so getting this far was not an anticipated or easy task. Nevertheless, Debbie and Barry Sittler, thank you for your love and care. I appreciate and love you both. There's really nothing more to be said.

My Indiana University of Pennsylvania compatriots and colleagues also deserve recognition. To my dissertation chairperson and committee members—Dr. Mary Beth Leidman, Dr. Zachary Stiegler, and Dr. Jay Start—thank you. And I say that in the most heartfelt and sincere manner possible. You collectively did so much to help me succeed during this process… and I simply cannot repay it. But I will do my best pay it forward as I progress through my career. I figure that's the best way to honor your contributions to my success. Additionally, I need to extend accolades to Dr. Chris Juengel, Dr. Lacey Fulton, and Dr. Vicky Ortiz. You each contributed in a huge way toward helping me to complete my data collection. I appreciate it immensely. Dr. Hillary Creely and Ms. Alexandra Lykissas, thank you for helping me to keep this dissertation on point for the School of Graduate Studies and Research. Dr. Luis Almeida, thank you for introducing me to the work of Dwyer & Lamberski, as well as helping me to conceptualize the original version of this study. Finally, I am grateful to Dr. Francis Dwyer, Professor Emeritus at Penn State University, for granting me permission to use his materials as part of this study; I hope you are pleased with my contribution to the field.

On a similar note, thank you to my California University of Pennsylvania colleagues and extended family. President Gerri Jones, Provost Bruce Barnhart, and Dean Douglas Hoover –

thank you for your encouragement and support throughout this process. You each contributed in your own way to making my success achievable. I am extremely grateful to each of you. I must also acknowledge Professor Bill Meloy, Professor Loring Prest, and Dr. Chris Harman for not only being great colleagues but also good friends. They were always ready to lend a compassionate ear. I appreciate it.

There are a few friends that I need to call out specifically because they've had a role in getting me to where I am today. Soon-to-be Dr. David P. Keppel, one of my best friends in the world, was available 24/7 to discuss and help me mull over anything at the drop of a hat… be it something as simple as deciding what cards to add to my *Magic: The Gathering* EDH deck… or as complex as helping me study for Comprehensive Exams or deciding which statistical methods to apply to my research. You cannot ask for a better friend and colleague. I also need to acknowledge my friend Dr. Douglas Cook. I met Doug early in my professional academic career and he taught me what it means to be a true member of the academy. He stressed the importance of not only being good at your work, but also taking service and scholarship seriously. He also taught me to keep my sense of humor and to trust my own instincts. Thank you, Doug. Also, Dr. Dana Hackley, thank you for proofreading, providing levity, and helping to bolster my resolve in some of my more stressful moments.

Dr. Milo J. Aukerman, biologist and singer for the band Descendents, is a personal hero of mine. He taught me, through his music and deeds, that you can live in more than one world while still being true to yourself. Doctor? Punk? Musician? Father? You can be all of those things. And as long as you maintain your passion… "thou shalt always go for greatness . . . thou shalt not have no idea . . . thou shalt not let anything deter you in your quest for All . . . if you believe, you shall achieve ALL! ALL! ALL!" (Stevenson & McCuistion, 1987, track 11). I had the opportunity to hang out backstage with Milo and the rest of the band, at a concert in Toronto, right after my wife and I got married. Our wedding song was "We"—which Aukerman wrote for his own wife—and the Descendents graciously dedicated the song to us that evening. Milo also gave me words of encouragement regarding the dissertation process. The whole experience was very helpful at that point in my life. And I am grateful to have had it.

Finally, I need to thank the game crafters and artists that helped to keep me sane throughout the doctoral process. The folks at NetherRealm Studios made studying tolerable with the 2011 reboot of *Mortal Kombat*. When I was studying for Comps with David, this was our go-to game for letting out frustration. Also, the fine folks at Harmonix Music Systems deserve credit for developing *Rock Band* – as that franchise is directly responsible for getting me interested in music within the context of educational games (as drumming in that game could be quite educational!) When I started my doctoral program I intended to study *Rock Band* and musical self-efficacy. The topic shifted somewhat, but it was really that game that led me to the research you'll find in this document. Dr. Jorge Cham, thank you, for your wonderful *Piled Higher and Deeper* comic strip. It's always funny – because it's always true. And it has brought me a lot of smiles. Thank you to musician Derek Vincent Smith (who performs under the stage name Pretty Lights). His album, *A Color Map of the Sun*, was a staple during my writing process. It's no exaggeration to say that I listened to it roughly 40 times. It's fantastic writing music. Moreover, an even bigger thanks to Ian Anderson and the band Jethro Tull for their album *Thick as a Brick*. I clocked in more than 100 listens to this album while working on my dissertation. It never fails to quickly put me in the writing zone… and I do not think I could have finished this without it.

I feel that I have saved the best for last. I mentioned at the beginning of this section that completing my dissertation, and by extension the PhD degree, is my finest accomplishment.

Professionally, and in some ways personally, that is true in the short term. But as I sit here writing, contemplating the process, my experiences, and how I finally got to this point… I am well aware that my first child—my daughter—is due to be born in the next few weeks. Last fall, after being stuck in the political doldrums of dissertation writing (which was in many ways beyond my immediate control), I found out that my wife and I would soon have a child. Suddenly, finishing the program took on a new meaning and much greater importance. With the help of my committee, and perseverance, I was able to get the process back on track. So here I am. PhinisheD. And ready for the next round of adventures. I'll end with this:

This dissertation is dedicated to Verity Lee Sittler.

I cannot wait to meet you in the next few weeks, my soon-to-be daughter. I love you already. Always know that. Thank you for giving me the extra incentive to persevere. I owe it all to you. *You* are my higher purpose. *You* are my greatest accomplishment.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO THE STUDY

Introduction

Digital educational games provide opportunities to engage learners in new and unique ways that may not be possible with other technologies. Doing so affords innovative methods of interaction, engagement, and assessment that may be leveraged both in and out of the classroom. And though using games in educational applications is not a novel concept, they are rarely used as instructional strategies in either K-12 or college classrooms. This is particularly true for digital games but may change in the immediate future. The New Media Consortium (NMC), a respected group of educational technology experts and practitioners responsible for the publication of the Horizon Report series lists games and gamification as two of the next major emerging educational technology trends to significantly affect both K-12 and all levels of higher education within the next two to three years (The New Media Consortium, n.d.; Johnson, Adams, & Cummins, 2012; Johnson, Adams Becker, Cummins, Estrada, Freeman, Ludgate, 2013a; Johnson, Adams Becker, Cummins, Estrada, Freeman, Ludgate, 2013b; Johnson, Adams Becker, Estrada, & Freeman, 2014; Johnson, Adams Becker, Estrada, & Freeman, 2015). Considering this, it is now a favorable time to focus on digital educational games as a method of instruction. Opportunities for research exist in examining not only whether games do or do not work in this context, but what, how, and why specific aspects of these media can be leveraged for improved learning outcomes.

It was once complicated, time consuming, and cost prohibitive for researchers and educators to develop digital games. The advent of new free or inexpensive game design technologies such as Game Salad, Torque 3D, Adventure Maker, and Unity have made it easier to make games by mitigating some of the barriers to entry that have kept people out of this space (e.g., coding skills). Design products like Thinking Worlds, OpenSim, SimWriter, and Raptivity were created specifically to help designers develop digital educational games and simulations at a fraction of the complexity, time, and expense that was once required. And products such as Appcelerator and Corona allow those with an interest or experience in coding to develop a single iteration of a game in an integrated development environment (IDE), and publish to multiple software ecosystems (e.g., Apple iOS, Android, Nook, and Kindle) at once. This saves significant time in design, development, and deployment. These new technologies not only provide opportunities for more people to get involved in digital educational game development and utilization, but also provide opportunities for researchers to study these games in a wider context.

Despite digital game-based learning being studied in a variety of ways, this medium is notably devoid of a specific and accepted model of development; no adequate or recognized framework for creating successful digital educational games exists (Westera, Nadolski, Hummel, & Wopereis, 2008; Yuxin, Williams, Prejean, & Richard, 2007). However, Amory (2007) points out that a variety of possible frameworks have been proposed over the years and more are in development. These frameworks differ in application as well as in context. Examples have been proposed by Mustaquim and Nyström (2012), Champsas, Leftheris, Tsiatsos, Terzidou, and Mavridis (2012), Annetta (2010), and Yu, Li, Zhou, Zhuo, and Liu (2009) among others. None have garnered mainstream success. Similarly, no accepted framework exists for how educational games should be integrated into a classroom, even though scholars have suggested different possibilities for how this could be accomplished (del Blanco, Torrente, Marchiori, Martinez-Ortiz, Moreno-Ger, & Fernández-Manjón, 2012). Further, research on digital educational games focuses on instructional strategies, but few studies address specific aspects of digital game assets or the elements of their interfaces (e.g., audio).

This study examines learning within a 3D gaming environment and the influence of audio and gender on achievement of different learning objectives. Information obtained via sound and music has an effect on how human beings construct knowledge (McAdams & Bigand, 1993), but there is a dearth of information available on this topic (particularly in the context of games) and the findings that are available are tangential to the subject. Additionally, some multimedia researchers (e.g. Bishop & Cates) have specifically indicated that there is a need for more research in this area. Bishop and Cates (2001) identified little research on sound in educational software and that trend has continued since they first acknowledged this necessity. Clark and Mayer (2008) state that there is evidence that digital educational game interfaces and elements (such as audio) affect learning outcomes in various ways, but they go on to say “However, we need more data to guide interface design. . . . We need evidence on how the interface of a simulation or game affects motivation and learning” (p. 375). It follows then that excellent digital education game design necessitates exploration of all aspects of interface design. This will help to determine not only how different parts of the interface do or do not affect learning outcomes but also how to best design interface elements to maximize learner benefits. This includes audio.

Statement of the Problem

The effects of nonspeech audio in instructional and interactive media have been scarcely studied in the past three decades (Bishop & Cates, 2001; Bishop, Amankwatia, & Cates, 2008; Gaver, 1989). Further, little research on video game music has been conducted (Zehnder & Lipscomb, 2006; Fassbender, Richards, Bilgin, Thompson, & Heiden, 2012). These two factors help to explain why almost no research exists on audio in digital educational games (Clark & Mayer, 2008; Bishop & Cates, 2008). This study will narrow this gap by addressing the following question: What are the effects of audio and gender within a 3D gaming environment on the achievement of different educational objectives?

Purpose of the Study

The purpose of the study is to investigate the effects of audio and gender within a 3D gaming environment on the achievement of different educational objectives. This study extends the theory of information processing (Miller, 1956) by investigating limited capacity (Lang, 2000) in the context of digital educational games.

Research Question

This study examines audio and gender within a 3D gaming environment as they affect students' learning of facts, concepts, procedures, and principles. These four items are collectively referred to as different educational objectives (Merrill, 1983; Ragan & Smith, 2004).

RQ: What are the effects of audio and gender within a 3D gaming environment on the achievement of different educational objectives?

Hypotheses

The hypotheses for this study are derived from the researcher's knowledge of the existing literature, the intended purpose of this study, and the research question. There is an expectation that students will display a significant difference in achievement on criterion tests between groups that receive no audio as compared to the groups that receive the default audio treatment or the self-selected audio treatment. Further, gender is not expected to have a significant effect on the results. These hypotheses were selected because existing research has found that a negative effect on student learning can occur when audio is utilized (Clark & Mayer, 2008) and yet other research has found that audio can have positive effects on student learning (Fassbender, Richards, Bilgin, Thompson, & Heiden, 2012). Further, when subjects have preferred listening conditions—and hear sound that they find interesting—their performance improves significantly (Nantais & Schellenberg, 1999). As such, the following null hypotheses will be tested.

H_o1: There will be no significant difference in achievement on criterion tests between students that receive the default audio treatment and those that receive no audio treatment.

H_o2: There will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive no audio treatment.

H_o3: There will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive the default audio treatment.

H_o4: There will be no significant difference in achievement on criterion tests between males and females.

Significance of the Study

This study is significant because it addresses a variety of theoretical, educational, practical, and industrial needs. It advances theory on information processing (Lang, 2000) by extending research on acquisition and storage of information (Atkinson & Shiffrin, 1968), as well as theory on digital educational game design and instructional software by informing potential design frameworks. Theory for multimedia design (Clark & Mayer, 2008) and development best practices is also advanced by helping to determine how audio affects the knowledge acquisition process.

Theoretical Justification

Information processing is the theoretical basis for research in this study. This concept addresses humans' ability to grasp information, encode it, store it for later use, and then retrieve it so that it may be utilized at a later time (Lang, 2000). This is an essential aspect of cognition and relates directly to learning and education. It is also a delicate process that is susceptible to interference. The mind has limited information processing capacity (Lang, 2000). This was first studied by Miller (1956) and has maintained relevance since that time. This limited information processing capability can interfere with any step in the process and therefore affect one's ability to properly grasp, encode, store, retrieve, and/or use information when needed (Lang, 2000).

The way in which people retrieve information is directly related to how it was encoded at the time it was processed (Clark & Mayer, 2008; Hannon & Craik, 2001; Tulving & Thomson, 1973).Therefore, for example, studying for a test while in a certain environment encodes the studied information in a way that may make it more difficult to retrieve and decode in a different environment (e.g., a classroom) later. It also means that using educational media could affect the learner's encoding and later retrieval of information (Craik & Lockhart, 1972). The dearth of research on background music and how it affects learners has resulted in some scholars calling for the effects of background music to be explored more deeply:

> In the modern western world, where music is readily available to everyone through radio, recordings, television and videos, and where recorded background music is played routinely in many public places, the need to understand the effects of music on our behavior and cognitive processing has become increasingly important. (Savan, 1999, p. 139)

How audio affects information processing could depend on how it is implemented. This is due to depth of processing. The most well remembered stimuli are those that were most deeply processed (Craik & Lockhart, 1972), as they undergo greater semantic analysis (Lockhart & Craik, 1978) and therefore produce greater opportunity for future retrieval of information. When things are not remembered well, it may be due to shallow processing – this is incomplete semantic analysis (Wang, Bastiaansen, Yufang, & Hagoort, 2011). Audio may affect depth of processing.

Pragmatic Justification

Digital educational games can be expensive and time consuming to produce, and there is no agreed-upon framework for how these games should be constructed (Westera, Nadolski, Hummel, & Wopereis, 2008; Yuxin, Williams, Prejean, & Richard, 2007). It is therefore difficult to maximize the medium's effectiveness as an instructional tool. Audio increases expense as well as adding to design and development time when producing a digital game—for example, the game *Scarface: The World is Yours* had an audio budget of $2,500,000 (Superannuation, 2014). Therefore, audio is not always viewed as a high priority aspect of digital educational game design and development, and it could be regarded as unnecessary and easily eliminated during the preproduction process. However, if audio is found to have a statistically significant effect on the achievement of different educational objectives, it would make sense from a practical standpoint to devote time and money to maximizing its potential in new games. Previous studies have found that using a gaming environment as a context had statistically significant effects on achievement of different learning objectives (Almeida, 2008; Downs, 2008; Cameron & Dwyer, 2005). However, since there is not much research on audio in this context, the digital educational game design industry has little foundation on which to base best practices. The same holds true for teachers who wish to create these games for their classrooms. Fassbender, Richards, Bilgin, Thompson, & Heiden (2012) specifically state that the current research on background music and learning is insufficient both in games and in other contexts. This is why the researcher is conducting this study. A clear answer regarding the efficacy of audio in digital educational game contexts does not exist and thus there is an opportunity to improve these types of media by exploring the topic.

Research Design Synopsis

A 2 X 3 factorial post-test only control experimental design (Cresswell, 2009) will be used for this study. There will be two independent variables: audio and gender. Audio will have three levels: none, default, and self-selected. Gender will have two levels: male and female. The dependent variable will be achievement scores on three criterion-referenced tests: identification, terminology, and comprehension as well as the composite scores for the aggregate of these tests. Student participants are the unit of analysis. A factorial two-way ANOVA will be utilized to calculate the results of the study. The alpha level will be set at .05 ($*p < .05$). Each group will play a short 3D digital educational game that addresses the human heart’s functionality and parts based on content created by Dwyer and Lamberski (1977). This will be followed by three criterion-referenced tests immediately being administered to the group in order to collect data.

Delimitations

This study has a narrow scope in order to preserve focus on its specific research question and the hypotheses that follow from it. It therefore necessitates a quantitative method of inquiry and a finite number of dependent and independent variables to be addressed. Music types used in the treatment are one such delimitation as throughout the pursuit of future study, a significant difference in outcomes could be found between two specific genres of music. It is also possible that using two different songs from the same genre could result in different effects. The scope of the study could become unwieldy if not for delimitations that are specific, clear, and rigid. These other possibilities for research or adjustments to the dependent and independent variables will be addressed in Chapter 5 as part of the discussion for future research.

Definition of Terms

3D Gaming Environment	A three-dimensional, interactive environment that works with naturally-occurring human perception by extending visual information in all spatial directions (Dalgarno & Lee, 2010).
Cognitivism	A group of cognitive theories that examine the learning process and knowledge acquisition as it relates to the reception, organization, storage, and retrieval of information (Ertmer & Newby, 2013).
Concepts	Items, which have names and/or attributes in common, that describe an idea or ideas (Merrill, 1987).
Different Educational Objectives / Different Learning Objectives	Facts, concepts, procedures, and principles (Merrill, 1987; Cameron & Dwyer, 2005).
Edutainment	Software that conveys simple concepts to learners using basic digital games (Egenfeldt-Nielsen, 2007).
Facts	Objects, symbols, and events that share a relationship based on various criteria (Merrill, 1987). Things that are demonstrably true.

Games / Digital Games / Digital Educational Games	Any electronic games used for the purpose of teaching and learning (de Freitas, 2006).
Principles	The nature of a relationship, which can be causal or correlational (Merrill, 1987). A fundamental basis for certain types of information.
Procedures	A set of operations always presented and completed in the same manner (Merrill, 1987)
Serious Games	Digital simulations that deal with serious issues instead of entertainment (Schollmeyer, 2006).
Simulations	Digital games that that replicate the rules of the real world (DiPietro, Ferdig, Boyer, & Black, 2007).
Student Achievement	In terms of a particular learning outcome, performance on tests and grades achieved in classes (Romney, 2003).

Organization of the Remainder of the Study

The following chapters will delve more deeply into the topics of this dissertation. Chapter 2 provides a rich review of the literature pertaining to information processing theory, audio in educational media, digital educational games and simulations, and the application of audio in digital educational games. Chapter 3 contains an in-depth look at the formal research design and how the actual study was conducted (e.g., information relating to the independent and dependent variables, population, measurement instruments, and procedures). Chapter 4 presents the data as collected, formal analysis of the study, and results. Chapter 5 discusses results, conclusions, limitations, and possible directions for further research.

Summary

As discussed in this chapter, research on audio and how it affects digital game-based learning is lacking. Given the expectation that game based learning and gamification will become important in both K-12 and higher education within the next three years, better understanding is crucial and imminently needed. This study will be beneficial on a theoretical and pragmatic level in aiding digital educational game design and development for better learning outcomes.

CHAPTER 2

LITERATURE REVIEW

Introduction

When looking at digital educational games and simulations there is inevitably a question of how to use them effectively for learning. This includes identifying how to best leverage the strengths of this medium, mitigate design and efficacy weaknesses, and produce the best possible educational outcomes for all learners. This study examines the use of audio in this educational context and how to best utilize it for achievement. This requires looking at not only past and current research in the field, but also at how information processing applies to current learning theory as this is a historically important subject within the field of cognition and Cognitivism (Piccinini & Scarantino, 2011). Simply, understanding how people learn entails understanding how people think. This consideration of how people learn allows instructional strategies and materials to be more effectively designed, selected, and utilized.

Instructional approaches affect information processing (Bowman, Frame, & Kennette, 2013) and as such, understanding learning requires a thorough comprehension of information processing theory. However, researchers do not always agree on the role that certain aspects of this theory play in enhancing the efficacy of digital educational media. This is particularly prevalent when dealing with audio in digital educational games and simulations. To wit, Clark & Mayer (2008) believe that audio negatively affects learning while Bishop & Cates (2001) do not. Though this disagreement is certainly not limited to audio, it is an area with limited existing research. Combining understanding of learning theory and effective digital interactive media could help lead to the production of more effective digital educational games.

Information Processing

Examining information processing requires investigating cognitivism and the major psychological ideas that preceded it. Historically, this means looking at behaviorism, a favored psychological paradigm of at least fifteen similar yet distinct approaches to the study and analysis of behavior (O'Donohue & Kitchener, 1999). Behaviorism reigned as the leading standard in psychology from the early 1900s until at least the 1950s (Moore, 2011; Watrin & Darwich, 2012; Glassman & Hadad, 2009). All varieties of behaviorism focused on behavior and what could be understood about it, whether it was John B. Watson's emphasis on observable behavior or B.F. Skinner's separation of behavior from the mental/internal thoughts of an organism (Moore, 2011). But behaviorists paid little attention to how people think and what effect it had on outward behavior.

How people think is often referred to as cognition (Brown & Green, 2011) and falls under the psychological purview of cognitivism and cognitive theory. Cognitivism originated in the 1960s through the work of psychologists Kohler and Tolman (Glassman & Hadad, 2009) and contrasts with behaviorism since the focus is not on outwardly viewable behaviors, but on the internal mental processes of the mind. Cognitivism views the thinking and learning process as analogous to how computers process information (Leonard, 2002) – this includes thinking and learning that take place both inside and outside of the learner (Gagné, 1970). This notion of learning as information processing is supported by Gagné (1974), as well as other researchers in the fields of communication and learning theory (e.g., Baddeley and Clark & Mayer).

All of the simultaneous processes and subprocesses that a person experiences while dealing with information combine to form the phenomenon of information processing (Lang, 2000). As a theoretical framework, information processing looks at a person's ability to comprehend information, encode it, store it for later use, and then retrieve it for utilization in the future (Lang, 2000). This theory deals with general process development in humans which includes, but is not limited to various forms of memory (Gelman, 1999). It has early footholds in the computer architecture work of John von Neumann. His ideas about computer input units, central processing units, memory units, and output units directly influenced, and are essential to all models of the information processing theory (Sloman, 1999).

Information processing theory presupposes access to three types of memory stores (Atkinson & Shiffrin, 1968), each of which allows the mind to hold information for different and increasingly longer periods of time. These are, from shortest memory storage timeframe to the longest memory storage timeframe (Figure 1): sensory memory, working memory, and long-term memory (Rey, 2011). Information is not strictly moved from one of these three memory stores as it goes through the system, but is essentially copied and shared from one to the next (Atkinson & Shiffrin, 1968).

Sensory memory involves initial contact between a stimulus or message and a person's sensory receptors, such as the eyes or ears, which allows them to engage with information (Eysenck as cited in Lang, 2000). The learner receives information and alters it in a way that it can be stored in memory if needed (Driscoll, 2012) and is selectively encoded. A person may ignore information or choose not to move it forward to the next phase of the process (St Clair-Thompson, Overton, & Botton, 2010). This initial sensory memory is known as the sensory registry (Atkinson & Shiffrin, 1968), and is ephemeral in nature, holding information for a fleeting moment of time before it is either passed on to working memory or deteriorates entirely.

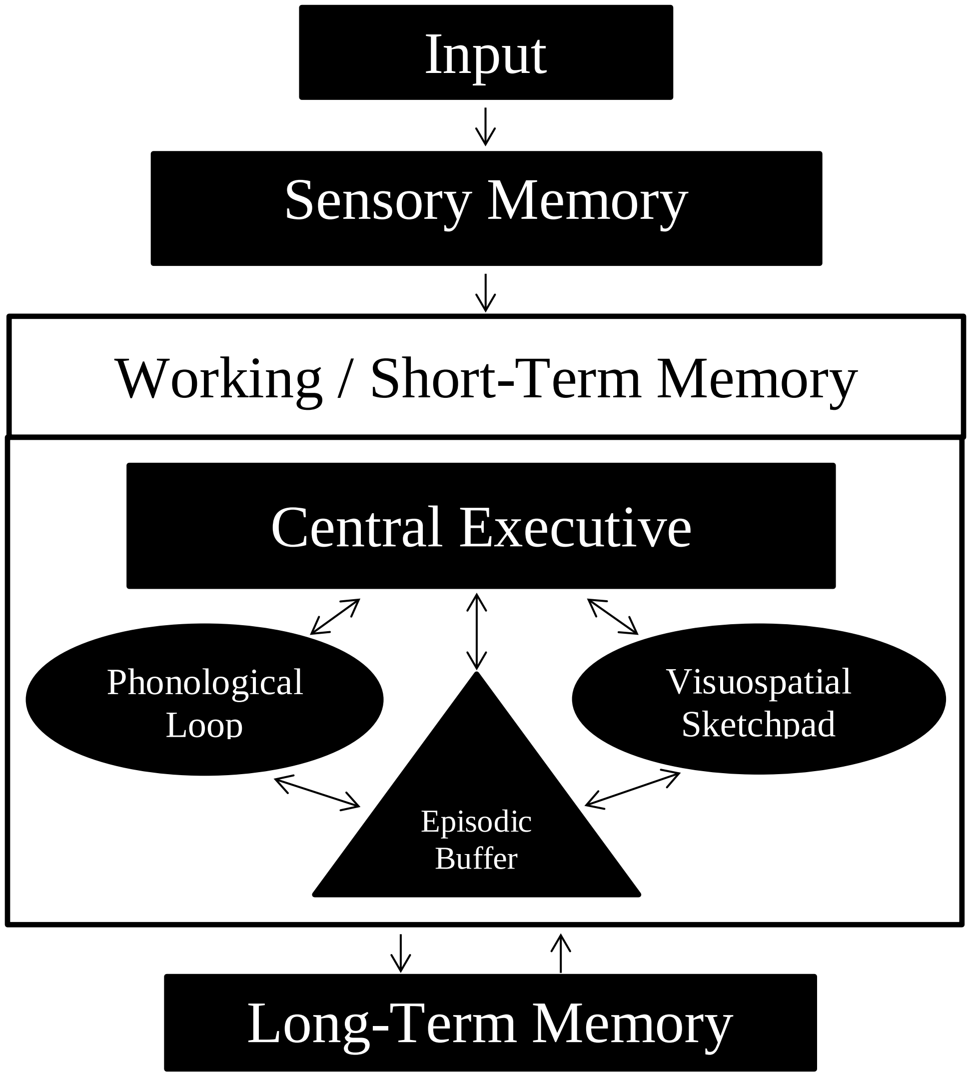


Figure 1. Information processing memory stores model. This Figure illustrates how input is passed from short-term to long-term memory (Atkinson & Shiffrin, 1968; Baddeley, 2010).

Also referred to as short-term memory (Driscoll, 2012), working memory is a system that allows the mind to engage in multifaceted cognitive tasks—including learning, reasoning, or comprehension—by providing a transitory workspace for the storage and manipulation of information (Baddeley, 1992; Baddeley, 2010). It also provides a workspace to connect the new information with that which already exists in long-term memory (Driscoll, 2012). The original model of working memory featured a single system (Atkinson & Shiffrin, 1968). A later model was produced by Baddeley and Hitch (1974) that differed significantly from the former in a variety of ways (Baddeley, 2010). The new model featured three components: the central executive system and two subordinate systems referred to as the phonological loop and visuo-spatial sketchpad (Baddeley & Hitch, 1974; Baddeley, 2010).

The central executive system controls attention and working memory (Baddeley, 2010). It is aided by both the phonological loop, which handles verbal-acoustic information, and the visuo-spatial sketchpad, which deals with visual information. Though this model existed for some time, it has since been updated to include a fourth component: the episodic buffer (Baddeley, 2000; Baddeley, 2010).The episodic buffer was added to this model in order to account for connecting information between working memory and long-term memory. It functions as a buffer zone for all components of working memory to come together, interact, and move on to long-term memory (Baddeley, 2000; St Clair-Thompson, Overton, & Botton, 2010; Baddeley, 2010). Working memory does not just involve information coming in from sensory memory and moving forward, it also receives and uses information from long-term memory (Atkinson & Shiffrin, 1968).

Researchers believe that working memory is limited in capacity (Atkinson & Shiffrin, 1968; Gagné, 1970; Baddeley, 2000; Clark & Mayer, 2008; St Clair-Thompson, Overton, & Botton, 2010), and that later recall can be enhanced by chunking information into smaller pieces so that working memory can process the information more efficiently. Starting with the work of Miller (1956), research has found that information held in working memory can be lost quickly. Upper limits suggested by Atkinson and Shiffrin (1968) suggest 30 seconds as the highest boundary of working memory (Gagné, 1970) though others suggest it may be as little as 20 seconds, especially if rehearsal is not involved in the memorization process (Peterson & Peterson, 1959).

Long-term memory is a permanent storage facility for all of our experiences and acquired knowledge and may have the strongest effect on human learning and performance compared to other memory stores (St Clair-Thompson, Overton, & Botton, 2010). Unlike both sensory and working memory stores, which quickly deteriorate, this is a permanent storage facility for information (Atkinson & Shiffrin, 1968). Various models have been proposed as to how long-term memory functions in information processing. A popular concept is that of schema theory and cognitive schemata. Schemata are mental constructs that help diminish workload in working memory (Schnotz & Kürschner, 2007). They are maintained in long-term memory and can be called upon to support information processing (Ericsson & Kintsch, 1995).

Deep processing of information creates better remembrance and recall (Craik & Lockhart, 1972), and can be influenced by how information was initially encoded when first processed (Clark & Mayer, 2008; Hannon & Craik, 2001; Tulving & Thomson, 1973). This deep processing of information is called semantic analysis (Lockhart & Craik, 1978). If semantic analysis is incomplete, information can be lost, inhibiting later retrieval (Wang, Bastiaansen, Yufang, & Hagoort, 2011). Information processing is a delicate system, particularly retrieving previously encoded information from long-term memory. Moreover, interference of any type can result in information being “destroyed, replaced, or lessened in value by subsequent information . . . information may never be destroyed but may become irretrievable, temporarily or permanently” (Atkinson & Shiffrin, 1968, p. 106).

Cognitive load is another factor that must be considered in information processing. Working memory can be overloaded when it holds and processes disparate information (Clark &

Mayer, 2008); while working memory retains information it also detracts from its ability to process additional information. While learning occurs, care must be utilized in limiting cognitive load and making sure that these restricted working memory resources are utilized to their greatest potential without overloading. Substantial use of interactivity in learning objects may unduly create cognitive overload and need to be reduced (Kalyuga, 2009). Learning theory suggests that humans have dual channels—auditory and visual—for processing information; this is typically referred to as dual coding theory (Clark & Paivio, 1991). Engaging both visual and auditory processes simultaneously is considered appropriate for learning (Clark & Mayer, 2008; French, 2006) and can reduce cognitive load. Therefore, careful planning and instructional design when producing educational media can mitigate issues and result in stronger learning effects.

Digital Educational Games and Simulations

Interest in using games, simulations, and virtual worlds in education has changed in conjunction with the shifting landscape of media during the past 60 years. This interest has also resulted in games' changing role in education. Game and simulation research started in the 1950s with a focus on business applications (Gros, 2007). As early computer games were produced in the late 1950s and early 1960s, more people expressed interest in studying them. This was due in part to the success of the first "modern" computer game called *Spacewar* (Lok, 2005) in 1962.

Game software dubbed "edutainment" emerged in the 1970s as more people had access to computers. This was an early attempt to marry learning theory with computer games (Egenfeldt-Nielsen, 2007) and continued unabated into the 1990s. These edutainment games, such as *Oregon Trail* or *Where in the World is Carmen Sandiego?* tended to be more successful commercially than educationally. Progressing and winning required little to no learning or knowledge acquisition. Typically, players were entertained and learned how to play the game well but got little else from the experience. As edutainment software proliferated in the 1980s, virtual worlds emerged concurrently as another computer-based platform for game players to use. These were typically multiuser dungeons (MUDs), and were also adopted for educational purposes over time (Peterson, 2006).

Interest in digital educational games and simulations increased and researchers began looking at results. However, findings were mixed. A 1992 meta-analysis of both digital and non-digital educational games and simulations empirical research examined 68 studies conducted between 1963 and 1991. Findings showed that usefulness of the game media depended highly on subject matter and that games were particularly helpful with math education (Randel, Morris, Wetzel, & Whitehill, 1992). Under the right circumstances, educational games did help with information retention. 39% of the studies showed games of any type to be a more successful teaching method versus traditional approaches. 56% of the studies found no difference in conventional teaching versus the use of games in education. 5% of the studies found that conventional teaching methods were more successful.

This inconsistency in digital educational game efficacy may be because no agreed-upon model or framework exists for how they should be created. Various structures have been suggested for use (Amory, 2007) in a variety of different contexts. However, none have been widely-adopted. One example would be Anneta's (2010) "I's model," which, though complex, is simpler than some others. The model includes a series of nested "I" statements which, from most central to all encompassing, include: identity, immersion, interactivity, increasing complexity, informed teaching, and instructional (p. 106). If all of these ideas are addressed appropriately, it is suggested that an effective serious game will be created. However, some researchers specifically state that no acceptable structure exists on which to base the creation of effective serious games (Westera, Nadolski, Hummel, & Wopereis, 2008; Yuxin, Williams, Prejean, & Richard, 2007). Similarly, classroom integration of this medium varies between classrooms despite a variety of integration techniques suggested by scholars (del Blanco, Torrente, Marchiori, Martinez-Ortiz, Moreno-Ger, & Fernández-Manjón, 2012); there is still no broadly-accepted framework for usage.

Digital educational games that address specific intellectual subjects emerged in the 2000s. These are typically called serious games (Corbit, 2005) although the term can be used interchangeably with others such as “games for change.” These serious games deal with issues like social change, health, and professional training (Schollmeyer, 2006) to educate or change the behavior of players. However, not all games that deal with training or education are considered serious games.

Children 18 and under are inundated with various media and usage has increased dramatically during the past few years. Children spent upwards of six and a half hours per day utilizing media of different types in 2008 (Annetta, 2008). A study by the Keiser Family Foundation (2010) found that usage by children aged 8 to 18 had increased to seven hours and thirty-eight minutes per day, seven days per week. The study points out that this equates to more hours than the average American adult spends working each week. This reflects more than one media type being consumed simultaneously due to multitasking (e.g. playing a video game while also listening to music). Multitasking accounts for approximately 29% overlap in time devoted to multiple media (p. 2).

Watching television and listening to music make up most of the time children 8 to 18 devote to media—approximately seven hours without accounting for multitasking (Keiser Family Foundation, 2010). Digital game usage accounts for approximately one hour and thirteen minutes of usage per day among this same age group (p. 2). Compared to other forms of interactive media that require active participation, children use digital games more than any other medium (Yuxin, Williams, Prejean, & Richard, 2007). Though just 58% of all Americans regularly play video games (Entertainment Software Association, 2013), research has found that this increases to 70% of college students (Martin & Ewing, 2008) and 97% of children aged 12 to 17 (Hoffman, 2009). Due to this prevalence, researchers and educators have explored opportunities to see how digital games can be utilized to benefit learners.

Almeida (2008; 2012), Downs (2008), and Cameron and Dwyer (2005) found that using digital game environments had a significant positive impact on achievement of different educational objectives. Empirical evidence shows that instructional games and simulations can also help learners to become better problem-solvers (Brown & Green, 2011). However, not all research on educational games has been positive or even conclusive. Some research suggests that games have either no affect – or a negative affect – on learning outcomes and therefore results regarding the efficacy of digital games are inconclusive (Clark & Mayer, 2008; Ke, 2008; Killi & Lainema, 2008; Prensky, 2008; Oliver & Carr, 2009).

Music and Cognition

Music directly impacts cognition and performance of mental tasks. Research traditionally focuses on participants listening to music composed by Mozart and is thus now known as the Mozart Effect (Chabris, 1999). This began with a seminal study in which subjects improved performance on abstract and spatial reasoning tests after listening to 10 minutes of Mozart's *Sonata for Two Pianos in D Major* (Rauscher, Shaw, & Ky, 1993). The improvement in performance lasted for a short time and diminished approximately 10 to 15 minutes after exposure (Rauscher, Show, & Ky, 1993; Rauscher, Shaw, & Ky; 1995). There was no improvement in performance after listening to silence or a relaxation tape. The authors suggest that “music lacking complexity or which is repetitive may interfere with, rather than enhance, abstract reasoning” (Rauscher, Shaw, & Ky, 1993, p. 611).

The same research team tried a separate experiment to see if daily exposure to Mozart’s music would show greater task performance improvement over the course of five days (Rauscher, Shaw, & Ky, 1995). The experimental group listening to Mozart underperformed (by less than a percentage point) as compared to the group exposed to silence on day one and tied performance with the group exposed to “mixed” audio. On day two there was a significant improvement in performance by the group exposed to Mozart: 62% for the Mozart group, 14% for the silence group, and 11% for the mixed group (p. 46). The Mozart group scored highest of the three groups on days three, four, and five. However, the silence group did not trail them by a significant amount and the mixed group scored lowest on these three days (p. 46). Regarding the results, the researchers stated that “the immediate improvement of the Mozart group’s scores was due to listening to the music, whereas the improvement of the silence group’s scores was probably the outcome of a learning curve” (p. 46).

Other scholars have extended the work of Rauscher, Shaw, and Ky. Nantais and Schellenberg (1999) were able to replicate results from earlier Mozart Effect studies. As with previous work, exposure to Mozart did significantly improve participant’s performance. However, the same effect was found to exist when subjects listened to a piece by Schubert. The authors suggest that this means any “enjoyable [Classical] pieces of music” would be successful (p. 371) for improving performance. They then conducted a second experiment in which subjects could be exposed to silence, music by Mozart, or a short story. Subjects exposed to Mozart or the short story both showed improvement compared to the silence control group. The researchers stated that participants experiencing their preferred condition performed better overall because it was more interesting to those subjects (p. 372) and there was a strong correlation between preference of listening condition and improved performance when compared to a non-preferred listening condition. The findings suggest any “engaging auditory stimuli” will result in significant performance improvement (p. 372).

The Mozart Effect may affect non-musicians differently than musicians; musicians may not experience any effects (Aheadi, Dixon, & Glover, 2010). 50 musicians and 50 non-musicians (p. 109) took a pre-test using 20 questions from a mental rotation test. This measures one’s ability to mentally “rotate” an object in various directions. They then listened to either “Mozart’s sonata for two pianos in D major (K448) for 15 minutes or an equal period of silence” (p. 109). Musicians were defined as having at least five years of formal musical training prior to reaching age 12. All participants were given a post-test of 20 mental rotation questions. Musicians’ performance remained stable regardless of exposure to Mozart’s music or silence; approximately 10 questions were answered correctly regardless of conditions. Non-musicians scored significantly higher on mental rotation tests after listening to Mozart; scoring less than 8 questions correctly before exposure and 10.5 after exposure (p. 111). The researchers conclude that this is due to activation of a specific hemisphere of the brain when listening to music. This is the same hemisphere that handles mental rotation tasks and spatial cognition (p. 112).

Despite Nantais and Schellenberg’s (1999) suggestion that any engaging auditory stimuli will result in performance improvement (p. 372), other researchers suggest that specific types of music may work better as an intervention with specific participants. A study by Hallam & Price (1998) found that 10 students with emotional or behavioral difficulties exhibited significantly less behavior problems while listening to calming music as a part of their normal school day:

17.3 rule-breaking instances with music and 21.1 without (p. 89). They also significantly improved performance on mathematics related tasks: average scores were 38.5 with music and 21.5 without (p. 89). These behavior and academic improvements were higher in students identified as hyperactive. Students identified as experiencing emotional or physical abuse showed improvement but at a lower level.

A similar study by Savan (1999) found that when 10 male students with special needs or emotional and behavioral problems listened to music composed by Mozart while working in a school setting, they experienced physiological changes including lower blood pressure, lower body temperature, and lower pulse rate (p. 142). They were also less disruptive, spent a greater amount of time on tasks, and had better work output than compared to the control group that received no music treatment. As Savan states, “the results clearly suggest that there are qualities present in certain Mozart orchestral compositions, which evoke changes in the pupils that directly affect the parameters of blood pressure, body temperature, pulse rate” (p. 143).

Hallam, Price, & Katsarou (2002) found that background music increased speed, but not accuracy, in a statistically significant way when students worked on mathematics problems. Those listening to silence had 80% accuracy versus 84% accuracy for students that listened to background music (p. 115). The same study indicated that in a separate but related experiment, calming music had a positive effect on memory task performance when compared to silence or aggressive music (p. 116). The researchers believe that this effect on performance is due to the effect of music on arousal and mood.

Not all literature supports the existence of the Mozart Effect; some research finds the effect is so small as to be insignificant and impractical. A meta-analysis of 20 Mozart Effect studies involving 714 subjects from 1993 until 1999 found that the cognitive improvement from listening to 10 minutes of Mozart's music before completing objectives resulted in enhancement of one specific task (Chabris, 1999). This result was not statistically significant (p. 826), and thus not useful. Steele, Bella, Peretz, Dunlop, Dawe, Humphrey, Shannon, Kirby Jr., and Olmstead (1999) attempted to replicate the original work by Rauscher, Shaw, and Ky, but found no statistically significant improvement among participants that listened to Mozart before completing assigned tasks. Contrary to previous work, they found that the group exposed to silence had the highest scores. Rauscher posits that these kinds of results could be tied to the fact that the person listening may just not enjoy music and therefore no effect would be presented (Spiegel, 2010).

Some researchers have chosen to examine specific aspects of music rather than the Mozart Effect directly. Ilie and Thompson (2011) found that music speed, pitch, and intensity can affect speed of information processing and object identification for task completion (p. 252). When asked to identify objects on a computer screen as quickly as possible, participants exposed to loud, fast, low-pitched music were able to successfully complete tasks faster. The same study showed that loud, slow, high-pitched music resulted in slower times. A separate study on auditory structural complexity found that higher levels of complexity created greater interest from participants (Potter & Choi, 2006, p. 408). Auditory structural complexity includes music as well as items like "sound effects, production effects, [and] voice changes" (p. 399). Participants experiencing higher levels of auditory complexity exhibited greater ability to freely recall data from memory that was introduced at the same time. As the authors summarize, "the impact of auditory structural complexity was statistically confirmed or marginally supported for 10 out of 13 dependent variables" (p. 412). Music was one aspect of this.

Audio in Educational Media

Bishop and Cates (2001) state that audio in educational media is typically addressed in one of three ways. It is used to narrate text, provide examples (e.g., sound recordings of speech or music), or to prompt learners when they have made a mistake. However, empirical research is almost non-existent and generally limited to narrated text (Bishop & Cates, 2001). Findings here are positive. There is scientific evidence that visuals presented with audio narration improve learning with greater significance as compared to visuals presented with written text (Clark & Mayer, 2012). This is because both channels (visual and auditory) are utilized. Carter (2012) believes this can be enhanced by following four design principles for instructional audio: selecting an appropriate narrative format for the audio, remaining cognizant of short term memory limitations when using spoken language, paying close attention to audio soundscapes (using foreground, background, and contextual sounds appropriately), and being deliberate in the use of sound and silence as needed (pp. 54-57).

Schwartz (2003) states that a message presented with audio cues or support will increase attention to the message due to arousal. This may improve the receiver's attitude towards the message while simultaneously overloading information processing and memory. Schwartz continues that this improved attitude toward the message is a hindrance. The overload causes interference with information retrieval and utilization. This defeats the purpose of using audio and is problematic when producing educational media. As such, Clark and Mayer (2008) argue that there is no theoretical justification for the addition of sound effects or music to educational media. They believe sound negatively impacts learning due to extraneous processing – information processing that is unrelated to the goals of the instructional system. They state "less is often more . . . leaner media can be more effective for learning" (Clark & Mayer, 2012, p.

318). Their studies have repeatedly shown that background music and environmental sounds reduced participant performance. Specifically, learners that did not receive an audio treatment had “between 61 to 149 percent better performance” in one study (Clark & Mayer, 2008, p.139). This adds credence to Gredler’s (2002) argument that background music could act as a distractor (and negatively affect performance).

Brown and Green (2011) disagree with Clark and Mayer’s assertion. They argue that audio, including background music and sound effects, can have a positive impact on learners’ information processing. Bishop and Cates (2001) propose that auditory cues can provide information not afforded by visuals. Further, Gaver (1989) argues that sound is an untapped modality and that computer interfaces will be positively augmented by the addition of nonspeech sounds to activities. These include icons being dragged across the screen or added to a trashcan, a task that is in limited practice in a variety of modern computer and mobile device operating systems. The disagreement between scholars regarding audio efficacy in educational media indicates that further investigation is required.

Gaver’s (1989) ideas are similar to those of Blattner, Sumikawa, and Greenberg (1989). They developed sound icons called Earcons. These work in tandem with graphic icons in a computer interface. The researchers concurred that sound effects have a place in interface design because audio is important to learning and memory. Mann (2008) also believes in the efficacy of audio in multimedia learning environments and proposed an extensive framework for engineering and using audio to help students focus on visual events in this context. Research has found that use of audio can increase attention and focus. Anderson & Levin (1976) used young children’s educational media—particularly *Sesame Street*—and found that sound effects and lively music gained the attention of viewers more successfully than visuals in the same programs. Children focused on those audio components as well as auditory change and adult voices, but responded highly to rhyming, repetition, and alliteration (p. 809).

Bishop, Amankwatia, and Cates (2008) studied sound in instructional software by conducting a mixed-method content analysis of sound usage in a variety of educational media. Twelve instructional software products were examined, each of which utilized narration. Five of these used narration exclusively, employing no other audio elements. Five of them used narration as the only sound source. Sound effects and music were both used in the remaining seven software packages, although usage was low. The instructional software using music most heavily accounted for just 5.27% of the total sound events utilized in the package (p. 476). The researchers also concluded that in the context of instructional media, both music and sound are not studied or discussed sufficiently.

A study using interactive media with story memory found that it did not have a strong effect on children's ability to successfully complete an assessment following the intervention (Daluz & Mapoy, 2011). However, students experiencing audio and video in unison during the intervention scored more highly than those that did not have this experience. The researchers note that results may have changed if used with a different type of interactive media, and suggest computer games as an opportunity for further study.

Schwartz (2005) studied the effects of animation and sound effects in educational media on attention and memory of multimedia messages. Effects were measured together and separately; results were mixed. A major finding was that the effects of sound on information retrieval via free recall were significant. Audio increased information retrieval better than when no audio was used. Information storage was assessed through cued recall and audio presented no significant difference in results. In sum, the study found that “sound effects aid information processing” (p. 130).

Audio and Music in Digital Educational Games

There is a lack of research on audio in digital educational games and what has been conducted mostly ignores music. This is not surprising as research on music in video games is rare (Zehnder & Lipscomb, 2006), and empirical research on music in educational games is even more scarce (Richards, Fassbender, Bilgin, & Thompson, 2008). Suggestions have been made for utilizing music in games (or music games) effectively. Hoffman (2009) states that music-based games like *Guitar Hero* and *Rock Band* can teach tenets of music training such as phrasing and rhythm (p. 21). Though logical, the statement is provided with no research to add veracity to the claim.

Few theoretical frameworks address how digital educational games might utilize nonspeech sound effectively. This is why Clark and Mayer (2008) are among those requesting that research be conducted in this area before specific recommendations are made for best practices. Fassbender, Richards, Bilgin, Thompson, & Heiden (2012) address this research gap by stating that “despite the apparent connection between music and cognitive tasks and the common use of music in computer games, the effect of background music on learning/memory has rarely been studied” (p. 491) [including in other educational media].

Yu, Li, Zhou, and Liu (2009) assert that since sound can be closely tied to visuals, sound should be explored with the addition of realistic multi-modal information presentation. An example of this is a visual of fire with matching fire sound effects. They believe this adds positively to the experience of the users. Making the environment more realistic and immersive when addressing something with real world application such as serious games requires that sound is lifelike for the same reason that graphics and physics need to be (Raghuvanshi, Lauterbach, Chandak, Manocha, & Lin, 2007). This could improve the experience for users. Gredler (2002) states that sound in digital educational games “may be distracting” (p. 572), particularly if paired and utilized with incorrect answers or information in a way that is more interesting to learners than correct answers or information. Despite these assertions, empirical evidence supporting them is scarce.

Audio and Music in Digital Educational Games and Effects on Achievement of Different Educational Objectives

Research on audio and music in digital education games is rare, particularly regarding effects on achievement of different educational objectives. Some researchers have found that digital educational games had a positive effect on achievement of different educational objectives (Almeida, 2012; Downs, 2008; Cameron & Dwyer, 2005). However, different educational objectives include facts, concepts, procedures, and principles (Merrill, 1987; Cameron & Dwyer, 2005) and a few studies have been conducted that examine the effects of audio and music in digital educational games on achievement on one of these factors.

Richards, Fassbender, Bilgin, & Thompson (2008) conducted an experiment using six variations of a 3D digital educational game, each using different background music (or lack thereof) as part of the treatments. Four distinct music selections were utilized, while one treatment utilized silence and another utilized a modified version of one of the original selections. These were soundtracks from the games *Oblivion*, *Baldur’s Gate*, *World of Warcraft*, and *Icewind Dale* (p. 236). One version of the game used this music with altered tempo and pitch. Another version of the game, administered to the control group, utilized no music stimuli. 72 undergraduate students participated in the study which involved receiving historical information about the Macquarie Lighthouse from an in-game avatar. Afterwards, they answered a post-experiment questionnaire that included 29 multiple-choice questions on historical facts about the lighthouse. Neither tempo nor pitch affected learning. Only the group that listened to the *Oblivion* soundtrack showed a statistically significant difference from the control group in terms of performance. The difference was positive, with participants that listened to the soundtrack from *Oblivion* answering 4.4 more questions correctly as compared to the control (p. 238). The researchers conjecture that this was because that soundtrack fit better with the theme of the game and created greater congruence. Participants self-reported that listening to music made it easier to concentrate (p.239) as compared to those that received no musical stimuli.

Fassbender, Richards, Bilgin, Thompson, & Heiden (2012) conducted an extension of this research built off of their previous findings. They used a modified version of the *Oblivion* soundtrack (from their previous experiment), as that was the only musical stimuli that had an effect. They also elected to forgo an interactive element in the game environment to better control variables and reduce interference with the experience (p. 493). 48 participants (28 female and 20 male) were broken into two experimental groups that received historical facts about the Macquarie Lighthouse for instruction in a 3D gaming environment. The first group of 24 participants listened to music while receiving instruction, during the first half of instruction only. The second group of 24 participants listened to music, while receiving instruction, during the second half of instruction only. Two different display systems were utilized. Participants then answered 30 multiple choice questions regarding historical facts presented during the instruction. Findings were mixed and inconclusive. The researchers sum this up in stating that “the results from this study show that in some circumstances music has a significant influence on memory in a virtual environment and in others it does not" (p. 499). However, 65% of participants preferred receiving instruction with music as opposed to the 35% that preferred silence (p. 497).

Summary

This chapter explained that information processing is a complex cognitive procedure that can be helped, or hindered, by sound. Appropriate use of sound is then necessary in designing digital educational games. Properly used, sound may enhance encoding and recall of learned concepts. Improperly used, information may be shallowly processed and encoded in such a way that makes it difficult to retrieve when the learner needs it. Little research has been conducted on how music specifically integrates with this process in regard to digital educational games. This study will add to this body of knowledge and aid in future framework development for this medium. Lessons learned in this review of the literature will be applied in the following chapter.

CHAPTER 3

RESEARCH DESIGN

Introduction

This study seeks to determine the effects of audio and gender within a 3D gaming environment on the achievement of different educational objectives. The research design combines the framework created by Dwyer and Lamberski (1977) for addressing the human heart's functionality and parts, as adapted by Almeida (2008), with audio selections influenced by studies on the Mozart Effect (Rauscher, Shaw, & Ky, 1993). At the time of writing, to the best of the Primary Investigator's knowledge, no other study has been conducted that combines these research elements in this manner.

Ethical Issues and Institutional Review Board Approval

There were no ethical issues anticipated with this study and no known ethical breaches were expected. There were no known psychological, personal, or academic risks to subjects associated with this study. Subjects were offered opportunities to discontinue participation or withdraw from the project if any discomfort arose. Indiana University of Pennsylvania (IUP) Institutional Review Board (IRB) approval was sought and granted in order to ensure that the study met with all appropriate University and Federal guidelines. Participation was voluntary (see Appendix C) and all participants were provided with and signed the Informed Consent Form (see Appendix D) before they could take part in the study. Subjects' names were collected on the Informed Consent Form but not connected with the data in any way. Likewise, IUP student ID numbers and personal information was kept anonymous. Identifying information is divorced from all data. Participants were made aware that they may request from the Primary Investigator a final report on the findings of this study and contact information was provided to do so.

Subjects' data is available if requested; it will not be identifiable by person as it is an aggregate of all the information collected.

Only the Primary Investigator and the Dissertation Chair have access to the data collected for this study. Physical data (e.g. consent forms) are kept in a locked cabinet in the Primary Investigator's office at California University of Pennsylvania. Electronic information collected is stored on Indiana University of Pennsylvania servers (which meet Federal security regulations for this type of data). All data, both physical and electronic, will be stored for a period of three years in compliance with Federal regulations. After this time period has expired, the data will be destroyed.

Characteristics of Subject Population, Method of Selection, and Recruitment

All participants were undergraduate students studying Communications Media at Indiana University of Pennsylvania who were at least 18 years old at the time of the study. There were no gender based restrictions on this study so both males and females were represented. Class rank (freshman, sophomore, junior, senior) was not a determining factor in participation and members of all groups were represented. No known vulnerable subjects participated.

The Primary Investigator contacted four instructors in the Department of Communications Media at IUP to find those willing to allow their students to participate in the study. A time was arranged for the Primary Investigator to attend the instructors' classes and explain the study to students (in one case, due to a scheduling conflict, the instructor announced the study to their class instead of the Primary Investigator) as part of recruitment. Participation was voluntary. At each classroom visit, the Primary Investigator read and provided students with a copy of the Letter of Invitation (Appendix A) and the Volunteer Sign-Up Sheet (Appendix B) as a part of the recruitment process. All students were given the opportunity to have a Volunteer

Sign-Up Sheet (Appendix B) so that no one felt influenced by their peers' participation. Students electing not to participate turned in a blank form, or in some cases, kept the form but did not turn it in.

Multiple dates and times were provided for participation in order to increase convenience and likelihood of students contributing. Some students chose to attend during open lab times held during November 2014 – Monday through Friday, 10:00am to 5:00pm. Other students attended during class times at the discretion of three of the four instructors. Students that participated were provided with and asked to sign the Voluntary Consent Form (Appendix C) and two copies of the Informed Consent Form (Appendix D). One copy was retained for their own records.

Compensation, Potential Benefits, and Alternatives to Participation

Students that chose to participate were fed pizza and soda, immediately after they completed the assessment, in exchange for their efforts (which they said was more enticing than extra credit would have been). The personal benefit for participating in this study was that subjects will have a greater understanding of the human heart and 3D educational game environments. Those students from targeted classes, that chose not to participate, were allowed to stop by and have some food even if they did not take part in the study.

Sample Size and Power Analysis

Sample size was a concern as without having a sufficient sample, the chance of incurring Type I and Type II errors increases (Pallant, 2013). The goal was to have at least 30 participants per group (Control, T1, and T2) in order to avoid problems with abnormal distribution of scores (Pallant, 2013). However, to double-check this assumption, a power analysis was conducted using the software package R (www.r-project.org/). R determined that at least 25.52 subjects were needed per group in order to reach acceptable power. The suggested syntax (Yu & Yagle,

2013) used in R to determine the subjects needed per group was > pwr.t.test(n =NULL , d = .8, sig.level =.05, power=.8). This resulted in a suggestion that n = 25.52 subjects were needed in order to attain an effect size of at least .80, a significance level of .05, and a power of .80 (Cohen, 1992). Figure 2 displays a visual guide for the minimum number of subjects desired per group in relation to the broader population.

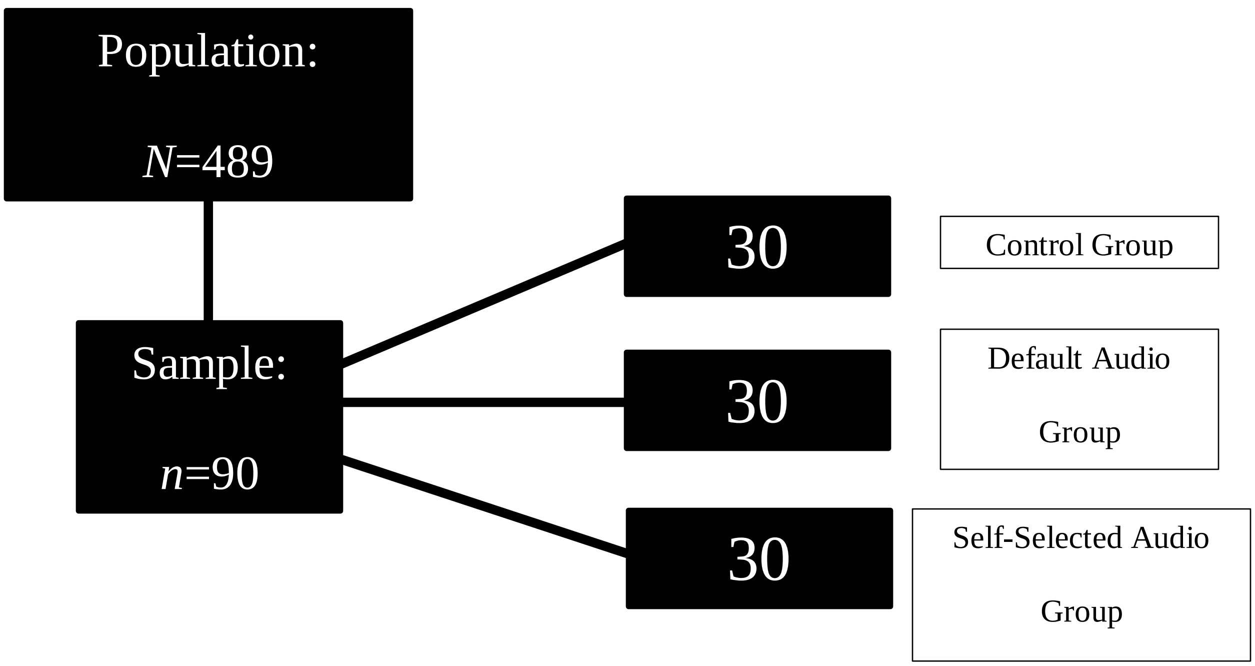


Figure 2. Visual guide for subjects per group. This Figure illustrates the minimum number of subjects desired for the Control, T1, and T2.

Development of Instructional Materials

The materials used for this study were modified from content created by Dwyer and Lamberski (1977) for addressing the human heart's functionality and parts. Though originally print based, similar to the work of Ausman (2008) and Almeida (2008), this study utilized digital versions of the resources and implemented them in a computer-based training (CBT) environment. Four major differences exist between the items used for this study and the original materials by Dwyer and Lamberski. First, the print materials included a visual depiction of the heart which included different parts labeled and identified as they were mentioned. This version of the test did not do so in order to reduce the number of variables included in the study. Second, the original criterion-referenced tests for assessing student achievement included four tests: identification, terminology, comprehension, and drawing. The drawing test was eliminated from this assessment in part because of the visual element being reduced. Third, after examining Cronbach's alpha coefficient for the full 60 question test, eight questions were eliminated, resulting in this version of the assessment employing 52 questions. Finally, the Almeida version of the instructional script and criterion tests had some slightly different wordings compared to the original material. This adapted version of the material was used for the present study.

Data collection was done with Qualtrics via a web-based survey and assessment that included all questions from the three criterion tests. The digital educational game used for this study was created in Caspian Learning's Thinking Worlds software (which is, at the time of this writing, no longer supported by Caspian). Programming for the game was completed by the Primary Investigator. Most of the assets used in the game came from pre-existing libraries of 3D models provided with the Thinking Worlds package (e.g. the classroom environment, desks, student non-player characters (NPCs), and Professor Hart – the instructor avatar NPC). Music files utilized for Treatment 1 were of Mozart's *Sonata for Two Pianos in D Major* under Fair Use as defined by Copyright Law of the United States (U.S. Copyright Office, 2011). This music was selected as it has been utilized extensively in Mozart Effect studies (Rauscher, Shaw, & Ky, 1993) that examine music, cognition, and learning. Screen captures of the instructional module can be found in this chapter (Figure 2, Figure 3, Figure 4, Figure 5). Structurally, the approach utilized in the instructional module is similar to that of Fassbender, Richards, Bilgin, Thompson, & Heiden (2012) in that interactivity was kept low and the instruction is guided by an NPC.

However, this was coincidental; the instructional module was designed and partially developed when that study was discovered.

Figure 3. Screen capture of NPC Professor Hart speaking to the Player. This Figure illustrates the opening moments of the instructional module.

Figure 4. Screen capture of the Player with decision tree for responding to NPC Professor Hart.

This Figure illustrates how player and NPC conversations take place in the instructional module.

Figure 5. Screen capture of the Player entering a classroom with various NPCs. This Figure illustrates how a sense of place and presence was created for the player in the instructional module.

Figure 6. Screen capture of the Player receiving instruction from the Professor Hart NPC. This Figure illustrates how instruction was provided to the player during the module.

Study Sites

The study was conducted at two facilities in Stouffer Hall at Indiana University of Pennsylvania's Main Campus. The first location, the Applied Media and Simulation Games Laboratory (room G12) had 15 PC-based student workstations available for simultaneous use. The second location, room G16, had more than 30 PC-based student workstations available for simultaneous use (and was utilized with larger classes). The Primary Investigator was present during the scheduled hours in order to ensure that directions were followed correctly. Equivalent computers and software were available at each site. Encore AE-06 disposable stereo headphones (http://bit.ly/1FYcVZK) were utilized for the experiment ; headphone frequency response was 20Hz – 20,000Hz (Encore Data Products Sales Representative, personal communication, January 30, 2015).

Measurement Instruments

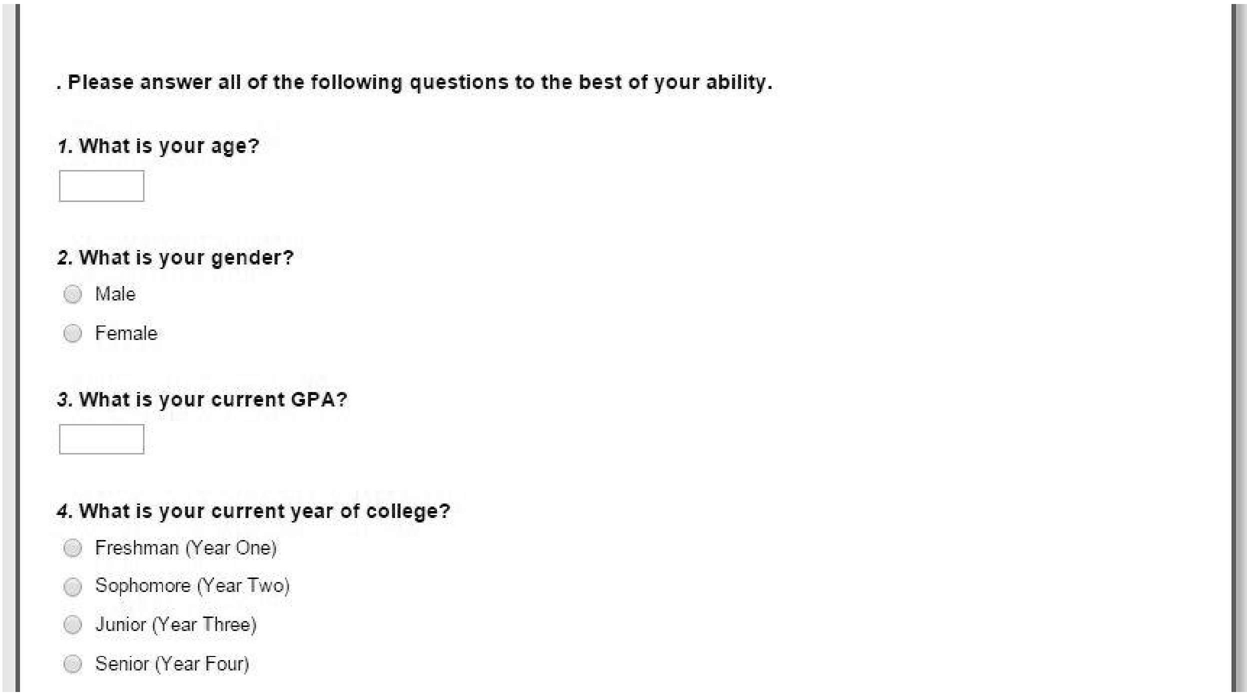
Three criterion-referenced tests were used to measure student achievement using three of the four heart content criterion tests developed by Dwyer & Lamberski (1977). Samples of the questions appear in Figure 8, Figure 9, and Figure 10. The full tests can be viewed in Appendix F. The original drawing test is not included as it was deemed inappropriate for this study. Additionally, demographic information for each participant was collected as part of the study. A sample of the demographic questions appear in Figure 7. This entailed 16 survey questions and can be viewed in Appendix E. The following is a brief overview of each test including details from Almeida (2008), Ausman (2008), and Lin (2006):

Identification Test. Normally a 20 question, 5 option, multiple choice test. The version utilized in this study was 18 questions with each being worth 1 point and a maximum score of 18 points. It measures students' ability to use visual cues to identify parts and names of the heart; factual level knowledge.

Terminology Test. Normally a 20 question, 5 option, multiple choice test. The version utilized in this study was 17 questions with each being worth 1 point and a maximum score of 17 points. It measures all areas as a pre-requisite to learning concepts, rules, and principles related to the heart.

Comprehension Test. Normally a 20 question, 4 option, multiple choice test. The version utilized in this study was 17 questions with each being worth 1point and a maximum score of 17 points. It measures procedures, principles, and understanding of heart function and parts: higher level information.

Composite Score. The aggregate score derived after completing each of the 3 tests. Normally 60 points, the version utilized in this study has a maximum score of 52 points. This score will be used to assess overall achievement for different educational objectives.



. Please answer all of the following questions to the best of your ability.

1. What is your age?

2. What is your gender?

Male

Female

3. What is your current GPA?

4. What is your current year of college?

Freshman (Year One)

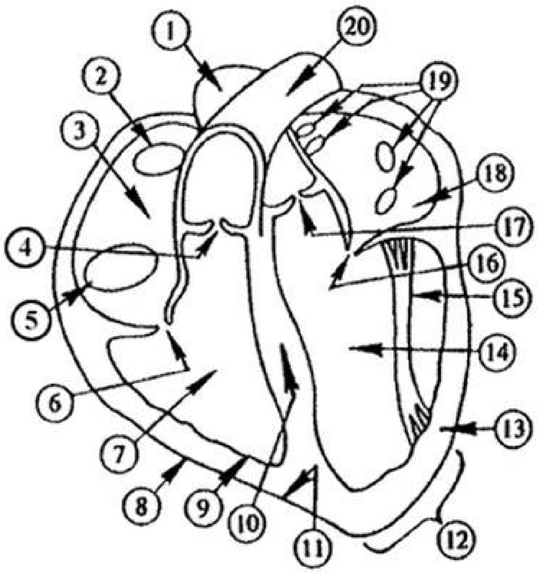
Sophomore (Year Two)

Junior (Year Three)

Senior (Year Four)

Figure 7. Screen capture of part of the Demographic Survey in Qualtrics. This Figure illustrates how questions were displayed during the module.

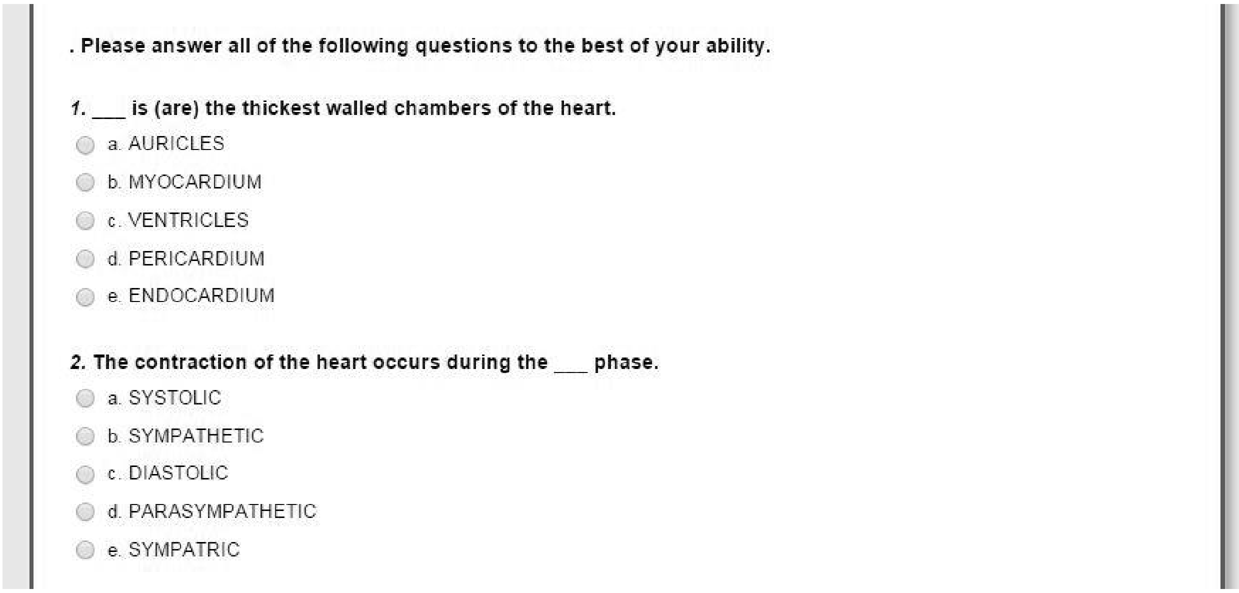
. **Please answer all of the following questions to the best of your ability.**



1. Arrow number one (1) points to the

- a. SEPTUM
- b. AORTA
- c. PULMONARY ARTERY
- d. PULMONARY VEIN
- e. NONE OF THESE

Figure 8. Screen capture of part of the Identification Test in Qualtrics. This Figure illustrates how questions were displayed during the module.



. Please answer all of the following questions to the best of your ability.

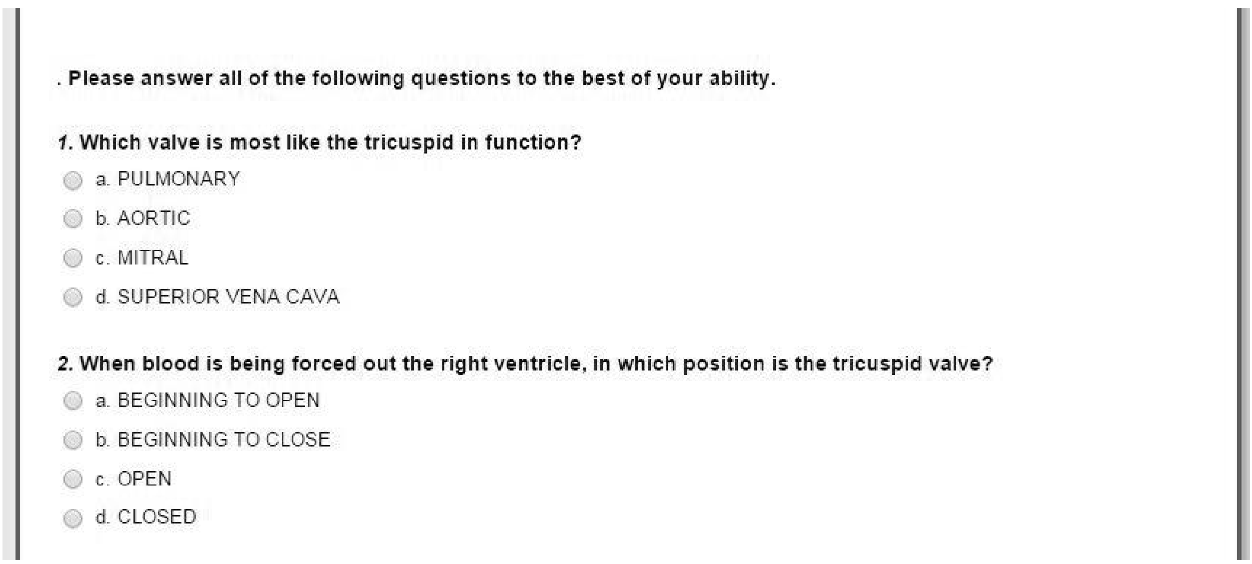
1. ___ is (are) the thickest walled chambers of the heart.

- a. AURICLES
- b. MYOCARDIUM
- c. VENTRICLES
- d. PERICARDIUM
- e. ENDOCARDIUM

2. The contraction of the heart occurs during the ___ phase.

- a. SYSTOLIC
- b. SYMPATHETIC
- c. DIASTOLIC
- d. PARASYMPATHETIC
- e. SYMPATRIC

Figure 9. Screen capture of part of the Terminology Test in Qualtrics. This Figure illustrates how questions were displayed during the module.



. Please answer all of the following questions to the best of your ability.

1. Which valve is most like the tricuspid in function?

- a. PULMONARY
- b. AORTIC
- c. MITRAL
- d. SUPERIOR VENA CAVA

2. When blood is being forced out the right ventricle, in which position is the tricuspid valve?

- a. BEGINNING TO OPEN
- b. BEGINNING TO CLOSE
- c. OPEN
- d. CLOSED

Figure 10. Screen capture of part of the Comprehension Test in Qualtrics. This Figure illustrates how questions were displayed during the module.

Instrument Reliability, Validity, and Adjustments

The "heart content" used in this experiment is based on a framework adapted by Almeida (2008) and created by Dwyer and Lamberski (1977) for addressing the human heart's functionality and parts. The original version of this training and assessment includes four tests. The fourth, a drawing test, is not utilized in this study as the participants would not be adequately prepared by the training to succeed in that test. As such, assessment for this experiment utilized the other three tests which initially totaled 60 items (20 per test).

After completing statistical analysis for each group, as can be seen in Table 1, the Cronbach alpha coefficient was $\alpha = .64$. Generally, this should be at least $\alpha = .70$ to indicate acceptable reliability (Pallant, 2013) and internal consistency. Since the assessment instrument is based on an existing scale it may have been appropriate to accept $\alpha = .64$ as indicated. However, since this is an adapted version of the test, poorly correlated inter-item values were assessed to determine if any questions should be dropped (Appendix H). 8 items showed negative inter-item correlation (questions 4 and 11 in the identification test; questions 1, 2, and 14 in the terminology test; and questions 10, 15, and 17 in the comprehension test). These questions were subsequently dropped from all scores which provided a new Cronbach alpha coefficient, as displayed in Table 2, of $\alpha = .71$ for the Composite test (table 3-2). The three criterion tests (Table 3) had lower Cronbach alpha coefficients: Identification $\alpha = .61$, Terminology $\alpha = .48$, and Comprehension $\alpha = .42$. These numbers, though low, also represent much smaller scales and so they may be acceptable (Pallant, 2013). Additionally, Validity for the purposes of this study was established via criterion-referenced validity as it is adapted from an existing scale (Reinard, 2006).

Table 1

Reliability Statistics Based on 60 Item Composite Test

Cronbach's α	Cronbach's α Based on Standardized Items	*N* of items
.636	.636	60

Note. Unacceptably low. Based on 60 questions.

Table 2

Reliability Statistics Based on 52 Item Composite Test

Cronbach's α	Cronbach's α Based on Standardized Items	*N* of items
.707	.710	52

Note. Acceptable. Based on 52 questions.

Table 3

Reliability Statistics for Identification, Terminology, and Comprehension Tests

Cronbach's α	Cronbach's α Based on Standardized Items	*N* of items	Test
.605 [a]	.602 [a]	18	Identification
.483 [b]	.485 [b]	17	Terminology
.416 [b]	.444 [b]	17	Comprehension

Note. [a] Acceptable based on 18 questions.
Note. [b] Acceptable based on 17 questions.

Methods, Experimental and Data Analysis Procedures

A 2 X 3 factorial post-test only control experimental design (Cresswell, 2009) was used for this study. There are two independent variables: audio and gender. The audio independent variable has three levels: none, default, and self-selected. The gender independent variable has two levels: male and female. Subjects were randomly assigned to one of three groups (Control, T1, and T2) by using a random number generator at a specialty website (www.random.org). Student participants are the unit of analysis. The study compares differences between and among these groups of subjects by utilizing achievement on criterion-referenced test scores as the dependent variable. These achievement scores are on three criterion-referenced tests—Identification, Terminology, and Comprehension—as well as a Composite score for those tests.

Cronbach's alpha coefficient was used to determine if the data collection instrument was reliable. After determining this, a one-way analysis of variance (ANOVA) was used to see if there was a statistically significant difference between the three groups. Since one was found, multiple factorial two-way ANOVA's were utilized to calculate the results of the study using IBM SPSS. The ANOVA was selected because it "allows you to simultaneously test for the effect of each of your independent variables on the dependent variable and also identifies any interaction effect" (Pallant, 2013, p. 275) and works especially well with two categorical independent variables combined with a continuous dependent variable.

In addition to the ANOVAs, Levene's Tests of Equality of Variance were used to check for equal variance between groups and both Tukey HSD and LSD post hoc tests were used to check for statistical significance. *A posteriori*, the researcher realized that a 3X3 factorial post-test only control design would have provided deeper analysis by including a three-way ANOVA that looked at gender, treatment, and another independent variable, whether or not the participant self-identified as a gamer or non-gamer. This analysis is included with the rest in Chapter 4. The alpha level was set at .05 ($*p <. 05$) for all tests. As of the Fall semester of 2015, Indiana University of Pennsylvania's Department of Communications Media, had a large population ($N = 489$) of undergraduate students taking coursework. Ideally, at least 30 subjects would have been assigned to each of the control and treatment groups ($n = 90$) to guarantee statistical significance. However, that number was not attainable and the total sample size was smaller ($n = 81$) with an equal number of participants ($n = 27$) in each group (Control, T1, and T2). This was still high enough to attain statistical significance given the results of the power analysis. Of the

81 participants, 100% completed the experiment. The following steps outline how the study took place:

1. Potential subjects were recruited by having the study explained by the Principal Investigator (in all but one instance) during a classroom visit. This was completed by reading the Letter of Invitation (see Appendix A).
2. Potential subjects were asked to fill out the Volunteer Sign-Up Sheet (Appendix B) to provide available times that they could participate in the study unless their professor was allowing class time for participation. This was returned to the Primary Investigator.
3. Potential subjects were asked to sign the Voluntary Consent Form (Appendix C) if they wanted to participate. This was returned to the Primary Investigator.
4. Subjects signed two copies of the Informed Consent Form (Appendix D) and gave one to the Primary Investigator while keeping the other for themselves.
5. Subjects reported to the study sites at their respective times.
6. Subjects were randomly assigned to one of three groups: Control, T1, or T2.
7. Subjects were given a DVD that contained their assigned game files which served as the intervention for this experiment. The game was run from their own disc.
8. Subjects were asked to select the computer of their choice at the study site. Each computer had headphones attached to it.
9. Subjects put on their headphones and then launched the game. Instructions for launching were provided via overhead projector and screen.
10. After starting the study, subjects completed a brief Demographic Survey (Appendix E) in-game.

11. Subjects received in-game instructions to test and keep volume level stable throughout the experiment. The default was 60% of maximum volume. Subjects were free to go as low as 40% of maximum volume and as high as 80% of maximum volume.
12. Within the game, subjects received instruction on the functionality and parts of the human heart, based on content created by Dwyer and Lamberski (1977), as adapted by Almeida (2008) (Appendix F). Each subject received the same instructional content and assessment instruments.
    a. The Control group listened to silence during the instruction; they received no audio.
    b. Treatment group 1 listened to Mozart's *Sonata for Two Pianos in D Major* during the instruction.
    c. Treatment group 2 listened to the music of their choice during the instruction using Spotify Web Player (play.spotify.com/) and were instructed how to do so in-game.
13. There was no time limit on completion; total time to complete instruction and criterion-referenced tests generally took less than 30 minutes.
14. After completing the in-game instruction, subjects immediately completed three criterion-referenced tests: Identification, Terminology, and Comprehension. Appendix F – Instructional Script / Instruments (Dwyer & Lamberski, 1977) as adapted by Almeida (2008).

15. After completing the three criterion-referenced tests, subjects removed their headphones and ejected the disc. The disc was returned to the Primary Investigator, subjects collected their food, and immediately left the study site in order to reduce background noise and disturbances to other test subjects.

Summary

All major aspects of conducting this study were discussed in this chapter. Ethical considerations and subject recruit plans were shared along with a thorough description of the instructional materials and assessment tools used in the experimental design. The procedures followed, data collected, and analysis resulted in a significant pool of data for discussion. This is explicated in Chapter 4.

CHAPTER 4

DATA AND ANALYSIS

Introduction

The purpose of this study was to examine how gender and music affect achievement in a 3D gaming environment when measured on three criterion-referenced tests for identification, terminology, and comprehension. The results of data analysis, tables, figures, and graphs created—as well as results of hypotheses tested in regard to the research question presented in Chapter 1—are all included in this chapter. All analysis was completed using a combination of the following software packages:

- Qualtrics (www.qualtrics.com). This is the official data collection tool of Indiana University of Pennsylvania. All demographic information and achievement test data were collected via the web using this software. Data were then exported to either Microsoft Excel or IBM SPSS depending on analysis need.
- Microsoft Excel (products.office.com/en-us/excel). This tool was used for simple data tracking, analysis, and line graph creation (e.g. long strings of data were stored here before being copied and pasted into data files in SPSS). In some instances collating data in Excel was easier than exporting directly to SPSS.
- IBM SPSS version 22 (www-01.ibm.com/software/analytics/spss/). This tool was used for the majority of data analysis and creation of histograms. All major statistical tests (e.g. ANOVA and Cronbach's Alpha) were completed using this software.

Further, all statistical techniques were double-checked to ensure that they followed appropriate assumptions for usage. Violations are noted, as applicable, in this chapter. The main statistical test utilized for this analysis was a two-way analysis of variance (ANOVA). The following assumptions must be met in order to utilize the test as per Pallant (2013, p. 213-215):

- Level of measurement – dependent variable(s) are measured at the interval or ratio level.
- Random sampling – participants were randomly sampled from the broader population.
- Independence of observations – each observed data set must be independent from others (i.e. separate subjects were used to collect data).
- Normal distribution – data will be distributed normally along a bell curve.
- Homogeneity of variance – populations have equal variance.

Participant Demographic Information

Table 4 displays demographic information for the subjects utilized in this study. This includes the group to which each subject was assigned (Control, T1, and T2), gender, mean age, mean GPA, current year in college, number of college classes taken that teach about the human heart (02 or 3+), and whether or not they identify themselves as a gamer. Broadly, 43 males took part in the study at a mean age of 21.82 and mean GPA of 3.18. 60% were seniors, none had taken more than 0-2 classes about the human heart, and 60% self-identified as gamers. 38 females took part in the study at a mean age of 21.19, and a mean GPA of 3.14. 61% of female participants were seniors, 8% had taken more than 3 classes about the human heart, and 16% self-identified as gamers.

Table 4

Demographic Information for All Subjects

Group	Gender	*n*	Age x̄	GPA x̄	Soph	Junior	Senior	0-2 [a]	3+ [b]	Gamer	Non-Gamer
Control	Male	14	21.64	3.15	2	7	5	14	0	10	4
	Female	13	21.08	3.17	1	3	9	13	0	1	12
	All	27	21.36	3.16	3	10	14	27	0	11	16
T1	Male	15	22.33	3.19	0	5	10	15	0	7	8
	Female	12	21.33	3.04	0	3	9	10	2	2	10
	All	27	21.83	3.12	0	8	19	25	2	9	18
T2	Male	14	21.5	3.19	0	3	11	14	0	9	5
	Female	13	21.15	3.21 [c]	4	4	5	12	1	3	10
	All	27	21.33	3.20 [c]	4	7	16	26	1	12	15
Total	Male	43	21.82	3.18	2	15	26	43	0	26	17
	Female	38	21.19	3.14 [c]	5	10	23	35	3	6	32
	All	81	21.51	3.16 [c]	7	25	49	78	3	32	49

[a] 0-2 classes taken that taught the participant about the human heart.
[b] 3+ classes taken that taught the participant about the human heart.
[c] One female participant in T2 did not indicate GPA. This datum was ignored.
Note. No Freshman took part in this study and so there is no data presented for that group.

Descriptive Statistics for Composite Scores

Table 5 displays the number of participants for the Control, T1, and T2. Each group had an equal number: $n = 27$. Additionally, it shows the mean and standard deviation of scores for the three tests—identification, terminology, and comprehension—as well as the composite scores for the three tests together. Both the Control group ($M = 14.59$, $SD = 5.37$) and T1 group ($M = 15.74$, $SD = 6.65$) outperformed the T2 group ($M = 11.67$, $SD = 4.04$) on the composite score for all three tests. T1 also outperformed the Control group.

Table 6 displays the number of male participants for the Control ($n = 14$), T1 ($n = 15$), and T2 ($n = 14$). It also shows the mean and standard deviation of scores for the three tests—identification, terminology, and comprehension—as well as the composite scores for the three tests together. Both the Control group ($M = 12.64$, $SD = 4.50$) and T1 group ($M = 14.27$, $SD =$

3.33) outperformed the T2 group (*M* = 11.79, *SD* = 5.01) on the composite score for all three tests. T1 also outperformed the Control group.

Table 7 displays the number of female participants for the Control (*n* = 13), T1 (*n* = 12), and T2 (*n* = 13). It also shows the mean and standard deviation of scores for the three tests—identification, terminology, and comprehension—as well as the composite scores for the three tests together. Both the Control group (*M* = 16.69, *SD* = 5.60) and T1 group (*M* = 17.58, *SD* = 9.16) outperformed the T2 group (*M* = 15.21, *SD* = 6.71) on the composite score for all three tests. T1 also outperformed the Control group. Tables 6 and 7 collectively show that female participants outperformed male participants in Control, T1, and overall composite scores. Males slightly outperformed females in T2.

Table 5

All Subjects' Scores by Test and Composite

		Identification Test[a]		Terminology Test[b]		Comprehension Test[b]		Composite Score[c]	
Group	*n*	x̄	σ	x̄	σ	x̄	σ	x̄	σ
Control	27	5.70	2.83	4.44	2.76	4.44	2.45	14.59	5.37
Treatment 1	27	5.59	3.30	5.30	3.16	4.85	2.93	15.74	6.65
Treatment 2	27	3.85	2.27	3.64	1.66	4.19	1.84	11.67	4.04
Total	81	5.05	2.92	4.46	2.67	4.49	2.43	14.00	5.66

[a] 18 Items
[b] 17 Items
[c] 52 Items

Table 6

Male Subjects' Scores by Test and Composite

		Identification Test[a]		Terminology Test[b]		Comprehension Test[b]		Composite Score[c]	
Group	*n*	x̄	σ	x̄	σ	x̄	σ	x̄	σ
Control	14	5.14	2.21	3.43	2.34	4.07	2.02	12.64	4.50
Treatment 1	15	5.20	3.10	4.33	2.02	4.73	2.76	14.27	3.33
Treatment 2	14	3.57	2.50	3.29	1.90	4.93	1.64	11.79	5.01
Total	43	4.65	2.69	3.70	2.10	4.58	2.18	12.93	4.34

[a] 18 Items
[b] 17 Items
[c] 52 Items

Table 7

Female Subjects' Scores by Test and Composite

		Identification Test[a]		Terminology Test[b]		Comprehension Test[b]		Composite Score[c]	
Group	*n*	x̄	σ	x̄	σ	x̄	σ	x̄	σ
Control	13	6.31	3.35	5.54	2.85	4.85	2.88	16.69	5.60
Treatment 1	12	6.08	3.60	6.50	3.94	5.00	3.25	17.58	9.16
Treatment 2	13	4.15	2.03	4.00	1.36	3.38	1.76	11.54	2.85
Total	38	5.50	3.13	5.32	3.00	4.49	2.72	15.21	6.71

[a] 18 Items
[b] 17 Items
[c] 52 Items

Composite Only One-Way ANOVA Results

A one-way between-groups analysis of variance was conducted to examine the effect of music on achievement as measured by the composite of three criterion-referenced tests. This was completed before the two-way analysis of variance in order to determine whether further testing was necessary. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 8 displays full statistics for this test. There was a statistically significant difference at the

*$p < .05$ level on scores for the three groups: $F(2, 78) = 4.00$, $p = .02$. The difference in mean scores was small (Table 9). The effect size, calculated using eta squared (*SS* between groups / total *SS*), was .09 and moderate to large. Levene's Test (Table 10) was not significant ($p = .192$) which indicates that the groups have equal variance (Pallant, 2013). Post-hoc comparisons using the Tukey HSD and LSD tests (Table 11) indicated that the mean score for the Control (M = 14.59, SD = 5.37) was not significantly different from either Treatment 1 (M = 15.75, SD = 6.65) or Treatment 2 (M = 11.67, SD = 4.04). Treatment 1 and Treatment 2 did differ significantly – with LSD indicating **$p < .01$. This indicated that further testing with a two-way analysis of variance was needed. The distribution curve for Figure 11 indicates that the data are not normally distributed; this is considered normal for social sciences research (Pallant, 2013).

Table 8

One-Way ANOVA Composite Scores Only

	SS	*df*	*MS*	*F*	*p*
Between Groups	238.296	2	119.148	3.996	.022
Within Groups	2325.704	78	29.817		
Total	2564.000	80			

Table 9

Descriptive Statistics for One-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test								
					95% Confidence Interval for *M*			
Group	*n*	$\bar{x}^a$	σ	*SE*	Lower Bound	Upper Bound	Minimum	Maximum
Control	27	14.59	5.37	1.03398	12.4672	16.7180	6.00	25.00
T1	27	15.75	6.65	1.28057	13.1085	18.3730	7.00	39.00
T2	27	11.67	4.04	.77717	10.0692	13.2642	6.00	23.00
Total	81	14.00	5.66	.62903	12.7482	15.2518	6.00	39.00

[a] 52 Items

Table 10

Levene's Test of Equality of Variances for One-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
1.688	2	78	.192

Table 11

Post Hoc Tests / Multiple Comparisons for One-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	Control	T1	-1.14815	1.48615	.721	-4.6990	2.4027
		T2	2.92593	1.48615	.127	-.6249	6.4767
	T1	Control	1.14815	1.48615	.721	-2.4027	4.6990
		T2	4.07407*	1.48615	.020	.5233	7.6249
	T2	Control	-2.92593	1.48615	.127	-6.4767	.6249
		T1	-4.07407*	1.48615	.020	-7.6249	-5.233
LSD	Control	T1	-1.14815	1.48615	.442	-4.1068	1.8106
		T2	2.92593	1.48615	.053	-.0328	5.8846
	T1	Control	1.14815	1.48615	.442	-1.8106	4.1068
		T2	4.07407**	1.48615	.008	1.1154	7.0328
	T2	Control	-2.92593	1.48615	.053	-5.8846	.0328
		T1	-4.07407**	1.48615	.008	-7.0328	-1.1154

* The mean difference is significant at the .05 level.
** The mean difference is significant at the .01 level.

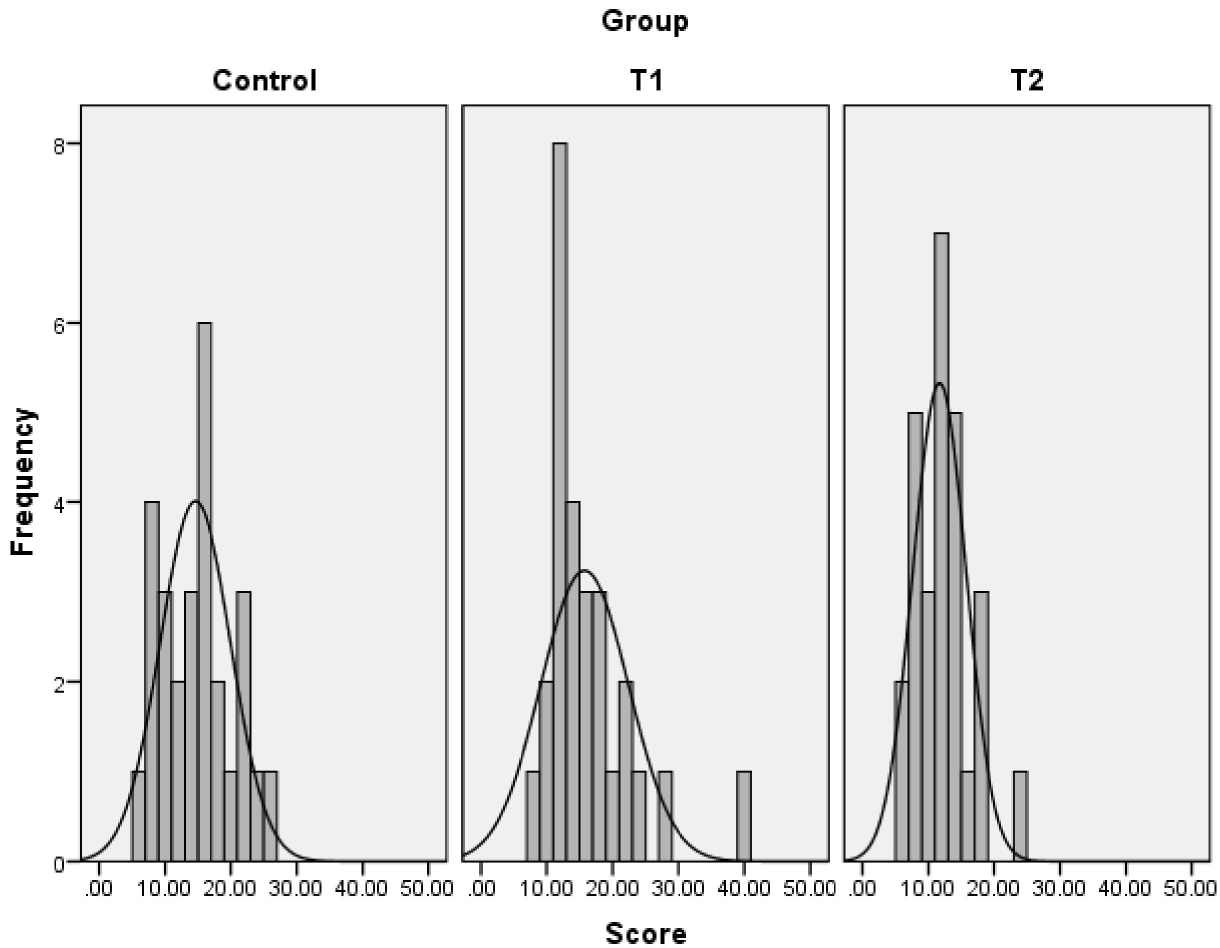


Figure 11. Histogram for One-Way ANOVA Composite scores only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Composite Only Two-Way ANOVA Results

A two-way between-groups analysis of variance was conducted to examine the effect of gender and music on achievement as measured by the composite of three criterion-referenced tests. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 12 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 75) = 1.25, p = .29. There was a statistically significant main effect for gender *F* (2, 75) = 3.97, *p* = .05 with a small-medium effect size (partial eta squared = .05). There was a statistically significant main effect for group *F* (2, 75) = 4.51, *p* = .01 with a small-medium effect size (partial eta squared = .11). The difference in mean scores was moderate (Table 13, Figure 12). Levene's Test (Table 14) was significant (*p* = .002), which indicates that the groups do not have equal variance; data were then transformed logarithmically in SPSS with the Lg10 transformation in order to better assess Levene's Test. After data transformation, Levene's Test (Table 15) was not significant (*p* = .06), and the groups were shown to have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 16) indicated that the mean score for the Control (*M* = 14.59, *SD* = 5.37) was not significantly different from Treatment 1 (*M* = 15.75, *SD* = 6.65). Tukey HSD indicates that Treatment 1 was significantly different from Treatment 2 (*M* = 11.67, *SD* = 4.04) **p* < .02. LSD indicates that Control was significantly different from Treatment 2 at **p* < .05 and Treatment 1 is significantly different from Treatment 2 at ***p* < .01. The distribution curve for Figure 13 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 12

Test of Between-Subjects Effects for Two-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	np^2	*Observed Power* [b]
Corrected Model	422.579[a]	5	84.516	2.960	.017	.167	.831
Intercept	15987.858	1	15987.858	559.950	.000	.882	1.000
Group	257.716	2	128.858	4.513	.014	.107	.754
Gender	113.449	1	113.449	3.973	.050	.050	.503
Group * Gender	71.161	2	35.581	1.246	.293	.032	.263
Error	2141.421	75					
Total	18440.000	81					
Corrected Total	2564.000	80					

[a] R^2 = .165 (Adjusted R^2 = .109)
[b] Computed using alpha = .05

Table 13

Descriptive Statistics for Two-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test				
Group	Gender	$\bar{x}^a$	σ	*n*
Control	Male	12.64	4.50	14
	Female	16.69	5.60	13
	Total	14.59	5.37	27
T1	Male	14.27	3.33	15
	Female	17.58	9.16	12
	Total	15.74	6.65	27
T2	Male	11.79	5.01	14
	Female	11.54	2.85	13
	Total	11.67	4.04	27
Total	Male	12.93	4.34	43
	Female	15.21	6.71	38
	Total	14.00	5.66	81

[a] 52 Items

Table 14

Levene's Test of Equality of Variances[a] for Two-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
4.284	5	75	.002

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 15

Levene's Test of Equality of Variances[a] for Two-Way ANOVA Composite Scores Only Post Transformation

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
2.284	5	75	.055

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 16

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Composite Scores Only

Dependent Variable: Scores on Composite Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Control	T1	-1.1481	1.45430	.711	-4.6255	2.3292
		T2	2.9259	1.45430	.116	-.5515	6.4033
	T1	Control	1.1481	1.45430	.711	-2.3292	4.6255
		T2	4.0741*	1.45430	.018	.5967	7.5515
	T2	Control	-2.9259	1.45430	.116	-6.4033	.5515
		T1	-4.0741*	1.45430	.018	-7.5515	-.5967
LSD	Control	T1	-1.1481	1.45430	.432	-4.0453	1.7490
		T2	2.9259*	1.45430	.048	.0288	5.8230
	T1	Control	1.1481	1.45430	.432	-1.7490	4.0453
		T2	4.0741**	1.45430	.006	1.1770	6.9712
	T2	Control	-2.9259*	1.45430	.048	-5.8230	-.0288
		T1	-4.0741**	1.45430	.006	-6.9712	-1.1770

Based on observed means.
The error term is *MSE* = 28.552
* The mean difference is significant at the .05 level.
** The mean difference is significant at the .01 level.

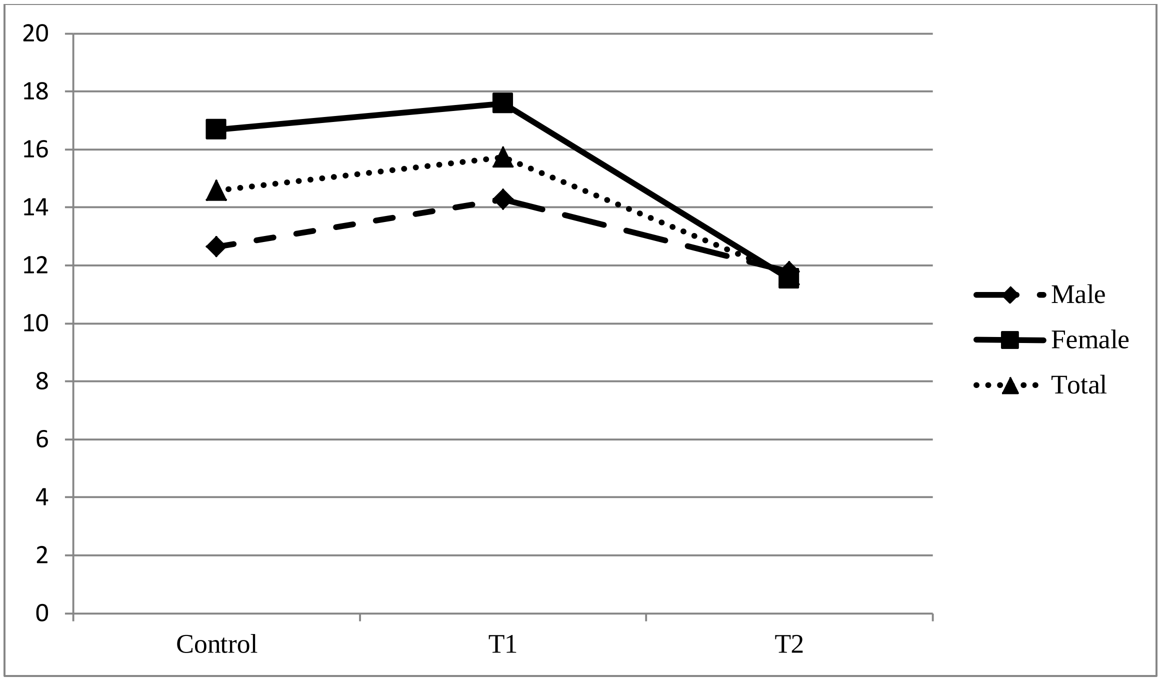


Figure 12. Means comparison for Two-Way ANOVA Composite scores only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

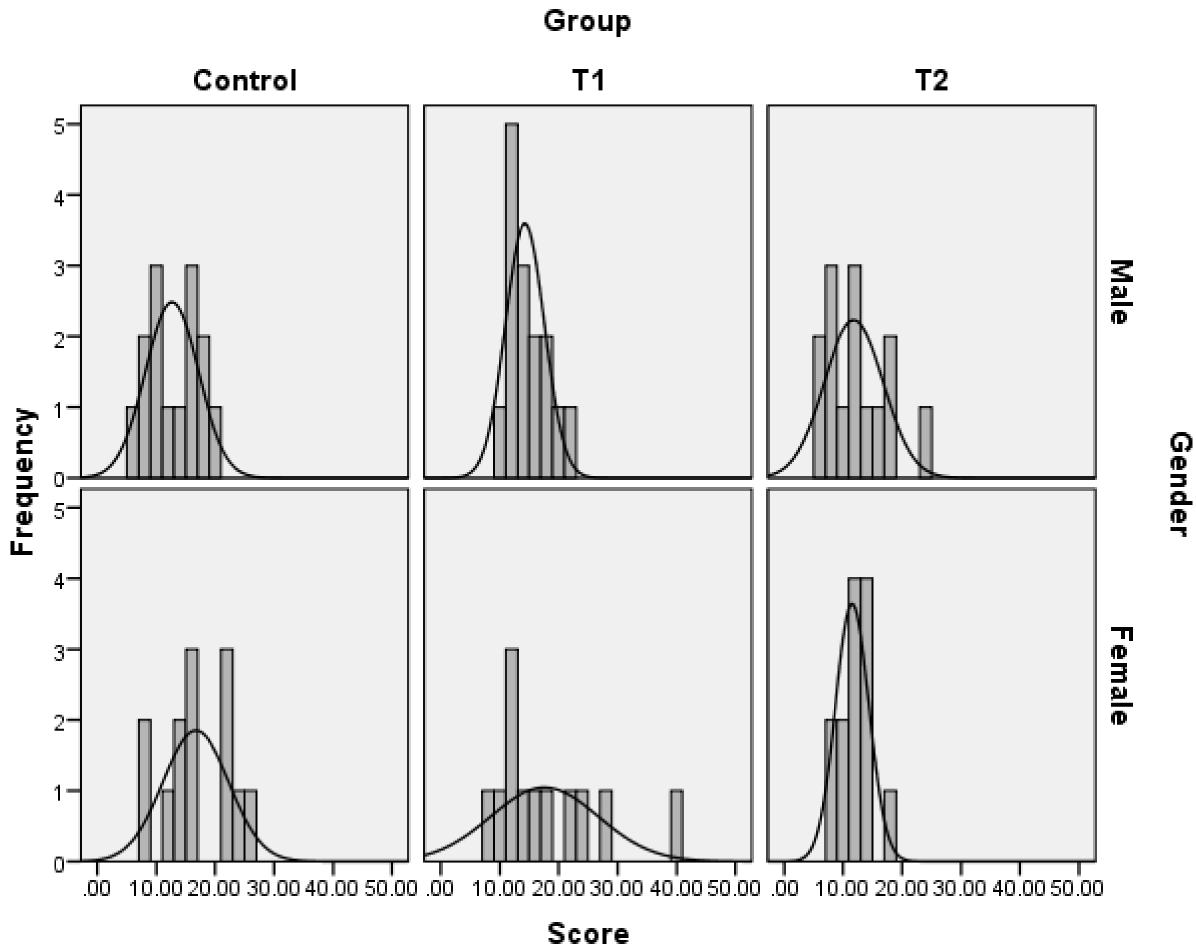


Figure 13. Histogram for Two-Way ANOVA Composite scores only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Identification Only Two-Way ANOVA Results

A two-way between-groups analysis of variance was conducted to examine the effect of gender and music on achievement as measured by the Identification criterion-referenced test. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 17 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 75) =.07, p = .93. There was not a statistically significant main effect for gender *F* (1, 75) = 1.91, *p* = .17 with a small effect size (partial eta squared = .03). There was a statistically significant main effect for group *F* (2, 75) = 3.68, *p* = .03 with a small-medium effect size (partial eta squared = .09). The difference in mean scores was small (Table 18, Figure 14). Levene's Test (Table 19) was not significant (*p* = .268) indicating that the groups have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 20) indicated that the mean score for the Control (*M* = 5.70, *SD* = 2.83) was not significantly different from Treatment 1 (*M* = 5.59, *SD* = 3.30). LSD indicates that the Control was significantly different from Treatment 2 (*M* = 3.85, *SD* = 2.27) **p* < .02. LSD indicates that Treatment 1 was significantly different from Treatment 2 at **p* < .03. The distribution curve for Figure 15 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 17

Test of Between-Subjects Effects For Two-Way ANOVA Identification Scores Only

Dependent Variable: Scores on Identification Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	*np*^{2}	*Observed Power* [b]
Corrected Model	74.881 [a]	5	14.976	1.851	.113	.110	.601
Intercept	2076.905	1	2076.905	256.653	.000	.774	1.000
Group	59.595	2	29.798	3.682	.030	.089	.660
Gender	15.491	1	15.491	1.914	.171	.025	.277
Group * Gender	1.144	2	.572	.071	.932	.002	.060
Error	606.921	75	8.092				
Total	2747.000	81					
Corrected Total	681.802	80					

[a] R^2 = .110 (Adjusted R^2 = .050)
[b] Computed using alpha = .05

Table 18

Descriptive Statistics for Two-Way ANOVA Identification Scores Only

Dependent Variable: Scores on Identification Test

Group	Gender	$\bar{x}^a$	σ	*n*
Control	Male	5.14	2.21	14
	Female	6.31	3.35	13
	Total	5.70	2.83	27
T1	Male	5.20	3.10	15
	Female	6.08	3.60	12
	Total	5.59	3.30	27
T2	Male	3.57	2.50	14
	Female	4.15	2.03	13
	Total	3.85	2.27	27
Total	Male	4.65	2.69	43
	Female	5.50	3.13	38
	Total	5.05	2.92	81

[a] 18 Items

Table 19

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Identification Scores Only*

Dependent Variable: Score on Identification Test

F	*df1*	*df2*	*p*
1.312	5	75	.268

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 20

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Identification Scores Only

Dependent Variable: Scores on Identification Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	Control	T1	.1111	.77423	.989	-1.7402	1.9624
		T2	1.8519*	.77423	.050	.0006	3.7031
	T1	Control	-.1111	.77423	.989	-1.9624	1.7402
		T2	1.7407	.77423	.070	-.1105	3.5920
	T2	Control	-1.8519*	.77423	.050	-3.7031	-.0006
		T1	-1.7407	.77423	.070	-3.5920	.1105
LSD	Control	T1	.1111	.77423	.886	-1.4312	1.6535
		T2	1.8519*	.77423	.019	.3095	3.3942
	T1	Control	-.1111	.77423	.886	-1.6535	1.4312
		T2	1.7407*	.77423	.027	.1984	3.2831
	T2	Control	-1.8519*	.77423	.019	-3.3942	-.3095
		T1	-1.7407	.77423	.027	-3.2831	-.1984

Based on observed means.
The error term is *MSE* = 8.092
* The mean difference is significant at the .05 level.

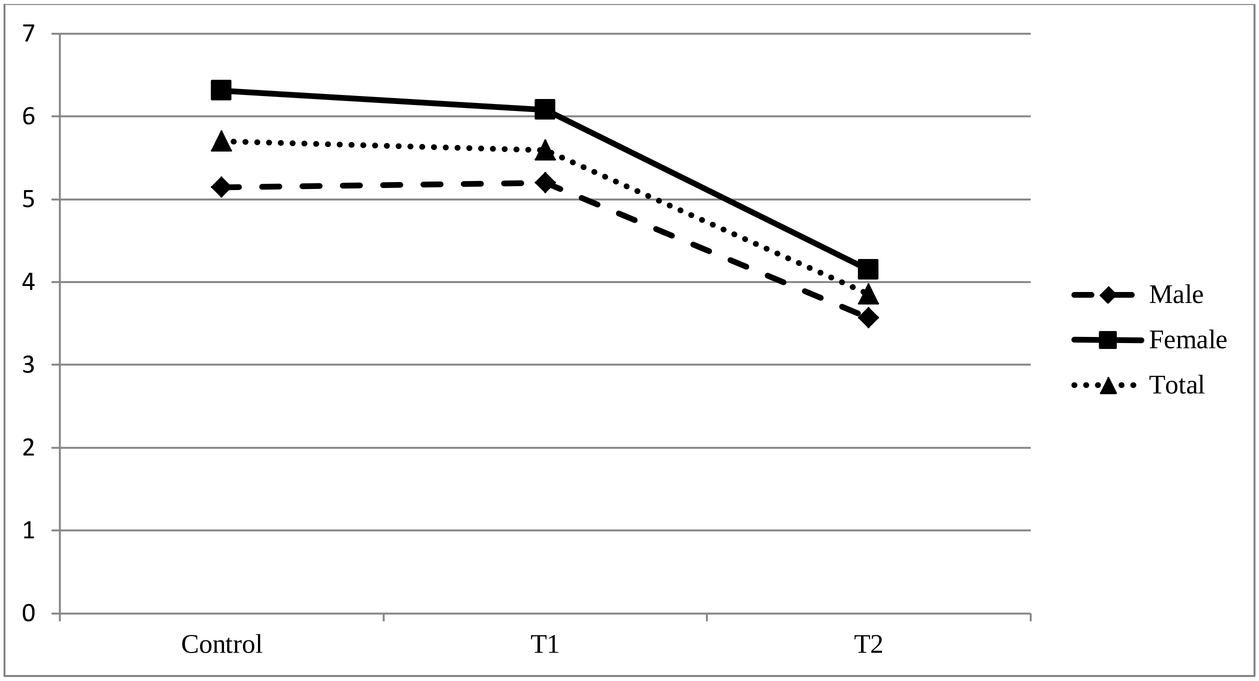


Figure 14. Means comparison for Two-Way ANOVA Identification scores only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

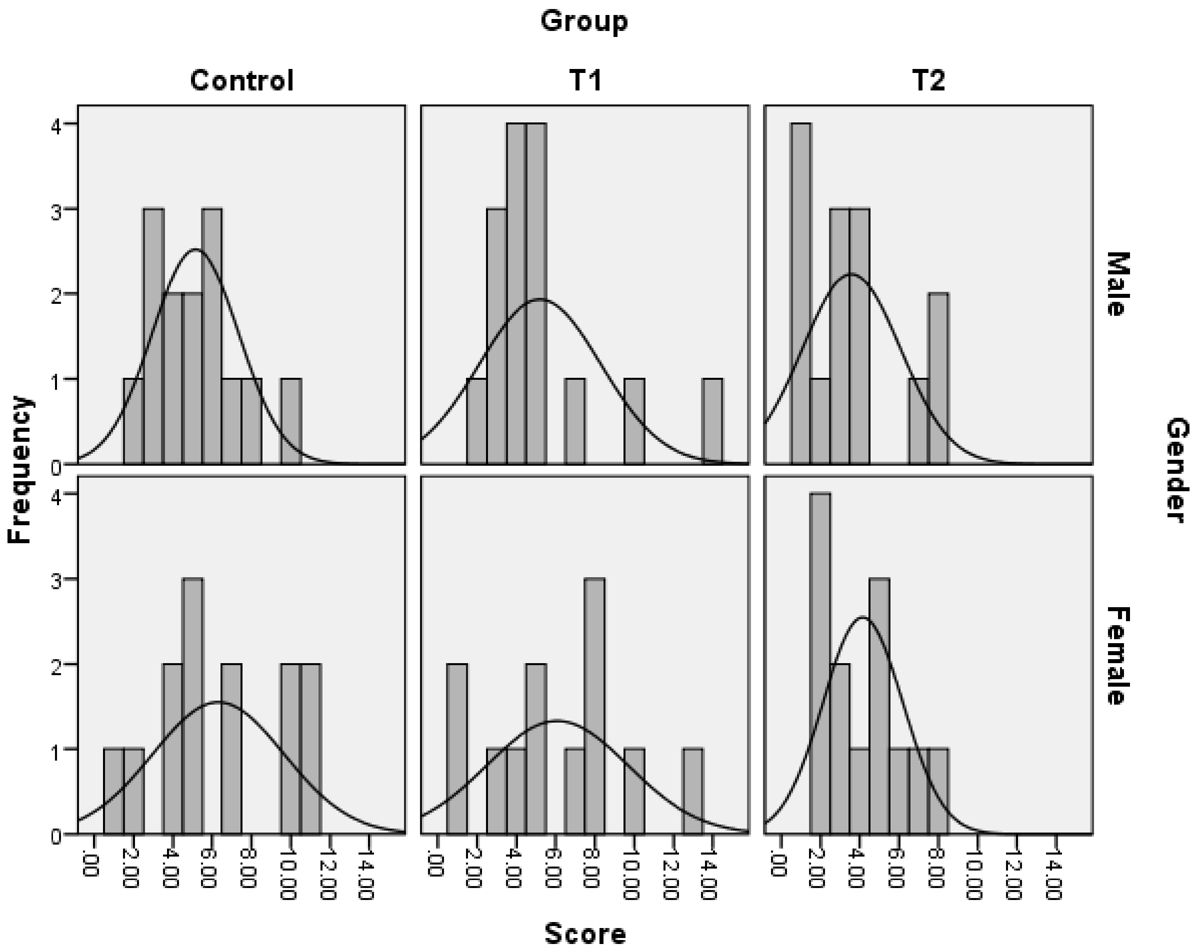


Figure 15. Histogram for Two-Way ANOVA Identification scores only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Terminology Only Two-Way ANOVA Results

A two-way between-groups analysis of variance was conducted to examine the effect of gender and music on achievement as measured by the Terminology criterion-referenced test. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 21 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 75) =.73, p = .49. There was a statistically significant main effect for gender *F* (1, 75) = 8.94, *p* = .01 with a small-medium effect size (partial eta squared = .11). There was a statistically significant main effect for group *F* (2, 75) = 3.38, *p* = .04 with a small-medium effect size (partial eta squared = .08). The difference in mean scores was small (Table 22, Figure 16). Levene's Test (Table 23) was significant (*p* = .003) which indicates that the groups do not have equal variance; data were then transformed logarithmically in SPSS with the Lg10 transformation in order to better assess Levene's Test. After data transformation, Levene's Test (Table 24) was not significant (*p* = .10), and the groups were shown to have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 25) indicated that the mean score for the Control (*M* = 4.44, *SD* = 2.76) was not significantly different from Treatment 1 (*M* = 5.30, *SD* = 3.16) or Treatment 2 (*M* = 3.63, *SD* = 1.69). LSD indicates that Treatment 1 was significantly different from Treatment 2 at $*p < .02$. The distribution curve for Figure 17 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 21

Test of Between-Subjects Effects For Two-Way ANOVA Terminology Scores Only

Dependent Variable: Scores on Terminology Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	np^2	*Observed Power* [b]
Corrected Model	102.249 [a]	5	20.450	3.278	.010	.179	.872
Intercept	1642.378	1	1642.378	263.286	.000	.778	1.000
Group	42.214	2	21.107	3.384	.039	.083	.621
Gender	55.761	1	55.761	8.939	.004	.106	.839
Group * Gender	9.104	2	4.552	.730	.485	.019	.169
Error	467.850	75	6.238				
Total	2179.000	81					
Corrected Total	570.099	80					

[a] $R^2 = .179$ (Adjusted $R^2 = .125$)
[b] Computed using alpha = .05

Table 22

Descriptive Statistics for Two-Way ANOVA Terminology Scores Only

Dependent Variable: Scores on Terminology Test				
Group	Gender	$\bar{x}^a$	σ	*n*
Control	Male	3.43	2.34	14
	Female	5.54	2.85	13
	Total	4.44	2.76	27
T1	Male	4.33	2.02	15
	Female	6.50	3.94	12
	Total	5.30	3.16	27
T2	Male	3.29	1.90	14
	Female	4.00	1.41	13
	Total	3.63	1.69	27
Total	Male	3.70	2.10	43
	Female	5.32	3.00	38
	Total	4.46	2.67	81

[a] 17 Items

Table 23

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Terminology Scores Only*

Dependent Variable: Scores on Terminology Test

F	*df1*	*df2*	*p*
4.041	5	75	.003

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 24

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Terminology Scores Only Post Transformation*

Dependent Variable: Scores on Terminology Test

F	*df1*	*df2*	*p*
1.959	5	75	.095

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 25

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Terminology Scores Only

Dependent Variable: Scores on Terminology Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	Control	T1	-.8519	.67976	.426	-2.4772	.7735
		T2	.8148	.67976	.458	-.8106	2.4402
	T1	Control	.8519	.67976	.426	-.7735	2.4772
		T2	1.6667*	.67976	.043	.0413	3.2921
	T2	Control	-.8148	.67976	.458	-2.4402	.8106
		T1	-1.6667*	.67976	.043	-3.2921	-.0413
LSD	Control	T1	-.8519	.67976	.214	-.5023	2.2060
		T2	.8148	.67976	.234	-.5393	2.1690
	T1	Control	.8519	.67976	.214	-.5023	2.2060
		T2	1.6667*	.67976	.017	.3125	3.0208
	T2	Control	-.8148	.67976	.234	-2.1690	.5393
		T1	-1.6667*	.67976	.017	-3.0208	-.3125

Based on observed means.
The error term is *MSE* = 6.238
* The mean difference is significant at the .05 level.

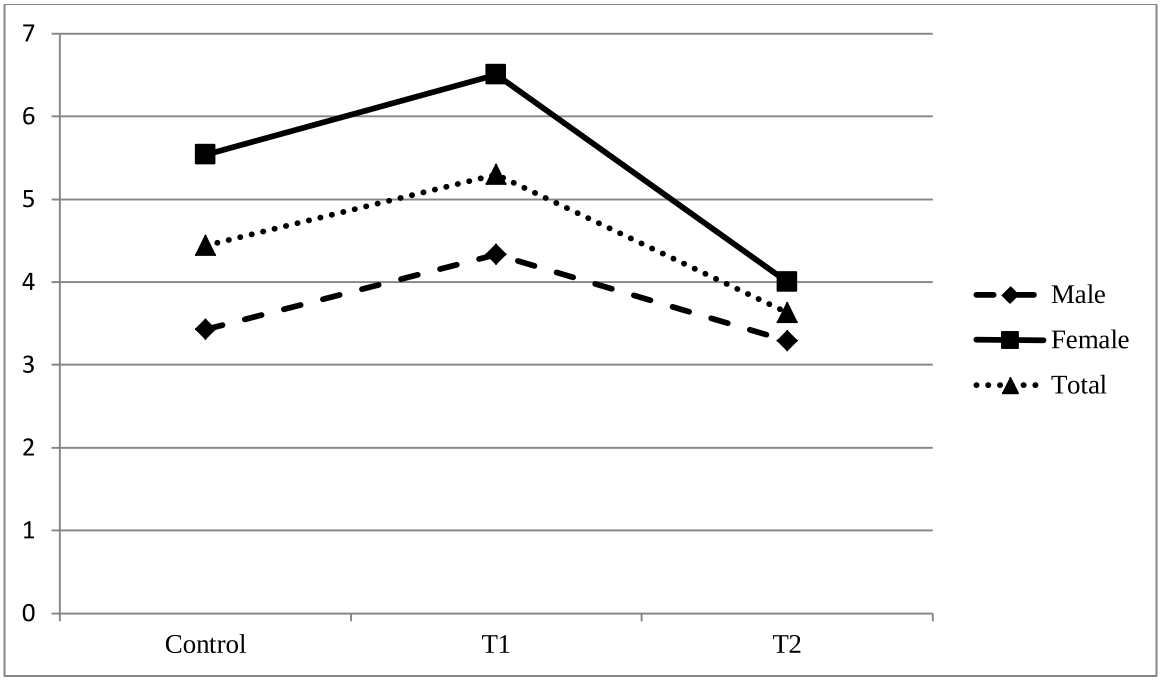


Figure 16. Means comparison for Two-Way ANOVA Terminology scores only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

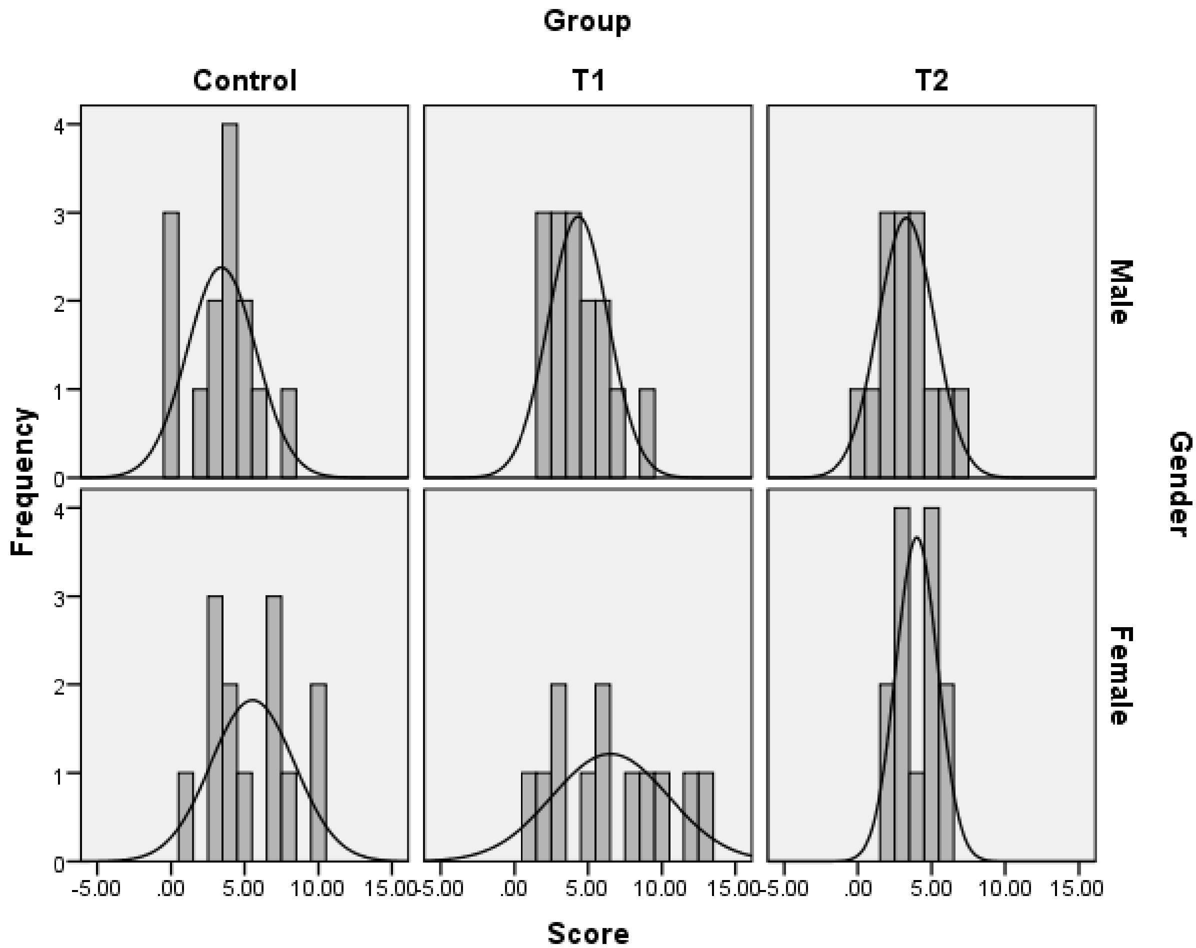


Figure 17. Histogram for Two-Way ANOVA Terminology scores only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Comprehension Only Two-Way ANOVA Results

A two-way between-groups analysis of variance was conducted to examine the effect of gender and music on achievement as measured by the Comprehension criterion-referenced test. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 26 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 75) = 1.68, p = .19. There was not a statistically significant main effect for gender *F* (1, 75) =.10, *p* = .76 with no effect size (partial eta squared = .00). There was not a statistically significant main effect for group *F* (2, 75) = 3.40, *p* = .57 with a small effect size (partial eta squared = .02). The difference in mean scores was small (Table 27, Figure 18). Levene's Test (Table 28) was not significant (*p* = .265), indicating that the groups have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 29) indicated that the mean score for the Control (*M* = 4.44, *SD* = 2.45) was not significantly different from Treatment 1 (*M* = 4.85, *SD* = 2.93) or Treatment 2 (*M* = 4.19, *SD* = 1.84). Further, both tests indicated no significant difference between Treatment 1 and Treatment 2. The distribution curve for Figure 19 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 26

Test of Between-Subjects Effects For Two-Way ANOVA Comprehension Scores Only

Dependent Variable: Scores on Comprehension Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	np^2	*Observed Power* [b]
Corrected Model	26.687 [a]	5	5.337	.894	.489	.056	.303
Intercept	1627.619	1	1627.619	272.749	.000	.784	1.000
Group	6.806	2	3.403	.570	.568	.015	.141
Gender	.565	1	.565	.095	.759	.001	.061
Group * Gender	20.012	2	10.006	1.677	.194	.043	.343
Error	447.560	75	5.967				
Total	2110.000	81					
Corrected Total	474.247	80					

[a] R^2 = .056 (Adjusted R^2 = .007)
[b] Computed using alpha = .05

Table 27

Descriptive Statistics for Two-Way ANOVA Comprehension Scores Only

Dependent Variable: Scores on Comprehension Test				
Group	Gender	$\bar{x}^a$	σ	*n*
Control	Male	4.07	2.18	14
	Female	4.85	2.88	13
	Total	4.44	2.45	27
T1	Male	4.73	2.76	15
	Female	5.00	3.25	12
	Total	4.85	2.93	27
T2	Male	4.93	1.64	14
	Female	3.38	1.76	13
	Total	4.19	1.84	27
Total	Male	4.58	2.18	43
	Female	4.39	2.72	38
	Total	4.49	2.43	81

[a] 17 Items

Table 28

Levene's Test of Equality of Variances[a] for Two-Way ANOVA Comprehension Scores Only

Dependent Variable: Scores on Comprehension Test

F	*df1*	*df2*	*p*
1.320	5	75	.265

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 29

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Comprehension Scores Only

Dependent Variable: Scores on Comprehension Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Control	T1	-.4074	.66486	.814	-1.9972	1.1823
		T2	.2593	.66486	.920	-1.3305	1.8490
	T1	Control	.4074	.66486	.814	-1.1823	1.9972
		T2	.6667	.66486	.577	-.9231	2.2564
	T2	Control	-.2593	.66486	.920	-1.18490	1.3305
		T1	-.6667	.66486	.577	-2.2564	.9231
LSD	Control	T1	-.4074	.66486	.542	-1.7319	.9171
		T2	.2593	.66486	.698	-1.0652	1.5837
	T1	Control	.4074	.66486	.542	-.9171	1.7319
		T2	.6667	.66486	.319	-.6578	1.9911
	T2	Control	-.2593	.66486	.698	-1.5837	1.0652
		T1	-.6667	.66486	.319	-1.9911	.6578

Based on observed means.
The error term is *MSE* = 5.967
Note. No mean difference significant at the .05 level.

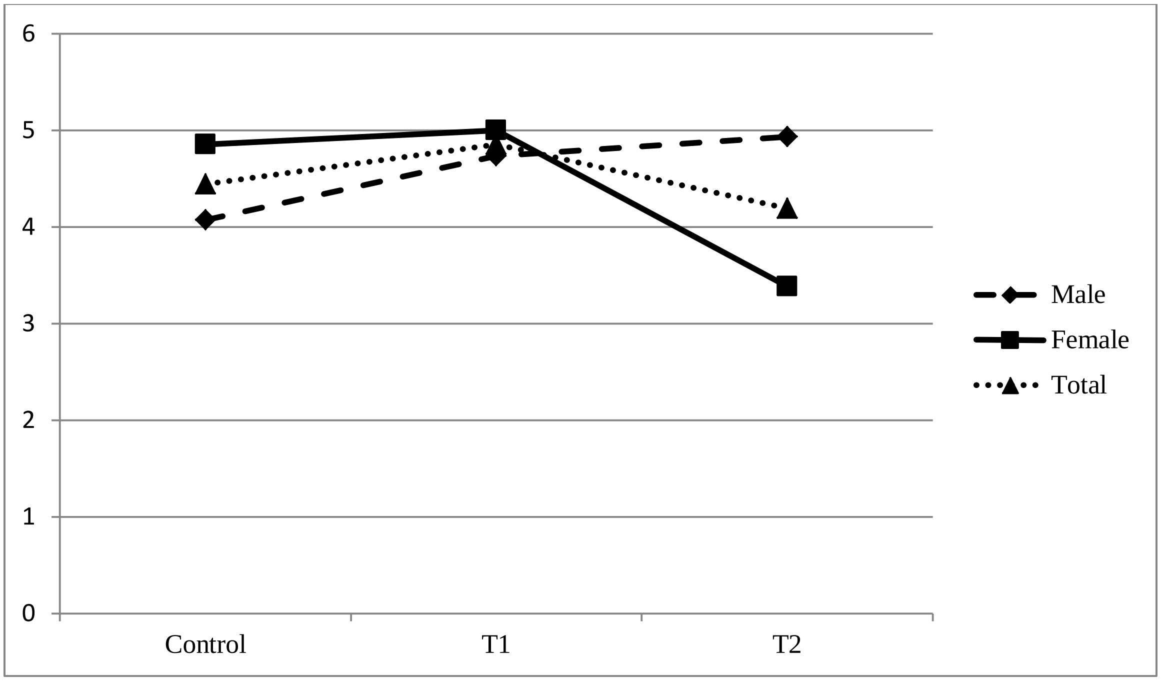


Figure 18. Means comparison for Two-Way ANOVA Comprehension scores only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

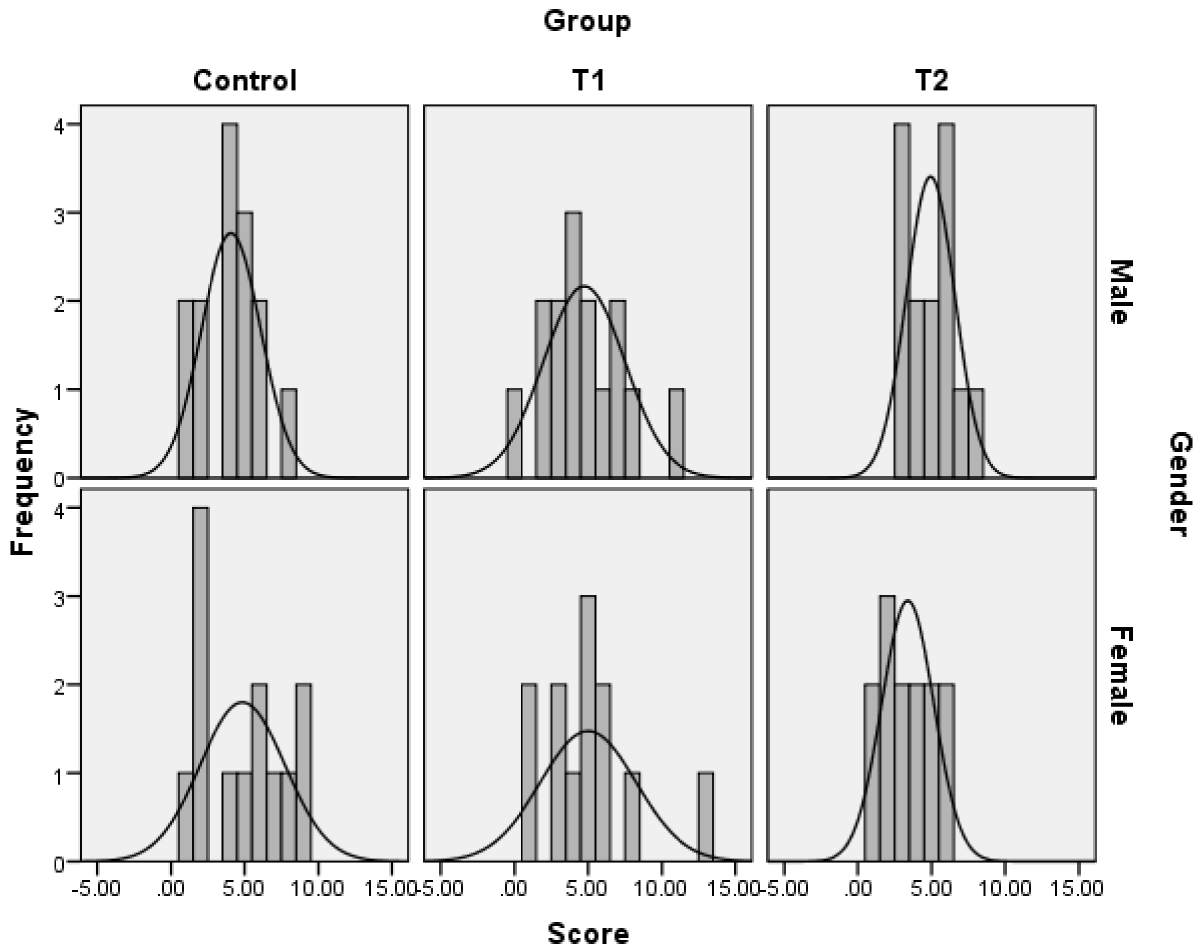


Figure 19. Histogram for Two-Way ANOVA Comprehension scores only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Composite With Classical Adjustment Only Two-Way ANOVA Results

A two-way between-groups analysis of variance was conducted to examine the effect of gender and music on achievement as measured by the composite of three criterion-referenced tests. This particular analysis was run to account for participants in Treatment 2 listening to classical music (which would make them more similar to participants in Treatment 1). Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 30 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 75) =.99, p = .38. There was a statistically significant main effect for gender F (1, 75) = 4.42, p = .04 with a small effect size (partial eta squared = .06). There was a statistically significant main effect for group F (2, 75) = 4.80, p = .01 with a small-medium effect size (partial eta squared = .11). The difference in mean scores was small (Table 31, Figure 20). Levene's Test (Table 32) was significant (p = .001) which indicates that the groups do not have equal variance; data were then transformed logarithmically in SPSS with the Lg10 transformation in order to better assess Levene's Test. After data transformation, Levene's Test (Table 33) was still significant (p = .03) and the groups were shown to not have equal variance. As such, Type I errors cannot be discounted and analysis must be done with caution. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 34) indicated that the mean score for the Control (M = 14.59, SD = 5.37) was not significantly different from Treatment 1 (M = 15.62, SD = 6.43). LSD indicated that Control was significantly different from Treatment 2 (M = 11.48, SD = 4.13) at *p < .04. LSD indicated that Treatment 1 was significantly different from Treatment 2 at **p < .01. The distribution curve for Figure 21 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 30

Test of Between-Subjects Effects for Two-Way ANOVA Composite Scores With Classical Adjustment Only

Dependent Variable: Scores on Composite Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	*np^2*	*Observed Power* [b]
Corrected Model	433.894[a]	5	86.779	3.055	.015	.169	.844
Intercept	15694.013	1	15694.013	552.578	.000	.880	1.000
Group	272.446	2	136.223	4.796	.011	.113	.781
Gender	125.430	1	125.430	4.416	.039	.056	.546
Group * Gender	56.412	2	28.206	.993	.375	.026	.217
Error	2130.106	75	28.401				
Total	18440.000	81					
Corrected Total	2564.000	80					

[a] R^2 = .169 (Adjusted R^2 = .114)
[b] Computed using alpha = .05

Table 31

Descriptive Statistics for Two-Way ANOVA Composite Scores With Classical Adjustment Only

Dependent Variable: Scores on Composite Test				
Group	Gender	$\bar{x}$[a]	σ	*n*
Control	Male	12.64	4.50	14
	Female	16.69	5.60	13
	Total	14.59	5.37	27
T1	Male	14.24	3.13	17
	Female	17.58	9.16	12
	Total	15.62	6.43	29
T2	Male	11.42	5.33	12
	Female	11.54	2.85	13
	Total	11.48	4.13	25
Total	Male	12.93	4.34	43
	Female	15.21	6.71	38
	Total	14.00	5.66	81

[a] 52 Items

Table 32

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Composite Scores With Classical Adjustment Only*

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
4.514	5	75	.001

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 33

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Composite Scores With Classical Adjustment Only Post Transformation*

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
2.689	5	75	.027

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 34

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Composite Scores With Classical Adjustment Only

Dependent Variable: Scores on Composite Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	Control	T1	-1.0281	1.42522	.752	-4.4360	2.3798
		T2	3.1126	1.47918	.096	-.4243	6.6495
	T1	Control	1.0281	1.42522	.752	-2.3798	4.4360
		T2	4.1407*	1.45445	.016	.6629	7.6184
	T2	Control	-3.1126	1.47918	.096	-6.6495	.4243
		T1	-4.1407*	1.47918	.016	-7.6184	-.6629
LSD	Control	T1	-1.0281	1.42522	.473	-1.8111	3.8673
		T2	3.1126*	1.47918	.039	.1659	6.0593
	T1	Control	1.0281	1.42522	.473	-1.8111	3.8673
		T2	4.1407**	1.45445	.006	1.2433	7.0381
	T2	Control	-3.1126*	1.47918	.039	-6.0593	-.1659
		T1	-4.1407**	1.45445	.006	-7.0381	-1.2433

Based on observed means.
The error term is *MSE* = 28.401
* The mean difference is significant at the .05 level.
** The mean difference is significant at the .01 level.

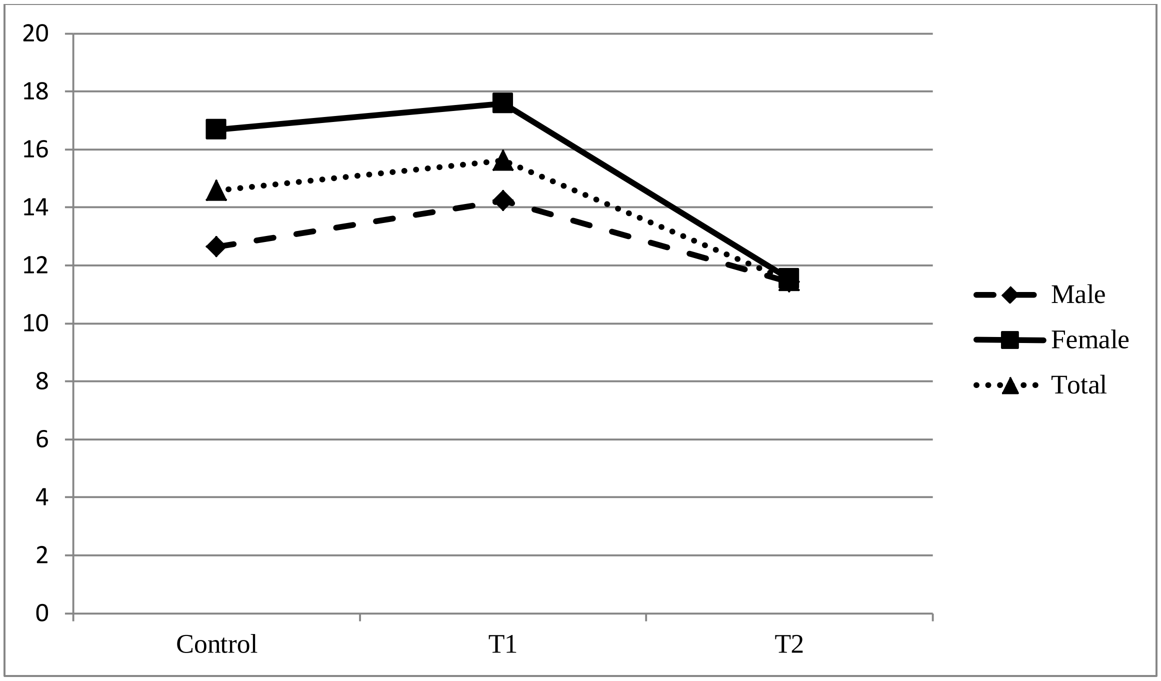


Figure 20. Means comparison for Two-Way ANOVA Composite scores with classical adjustment only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

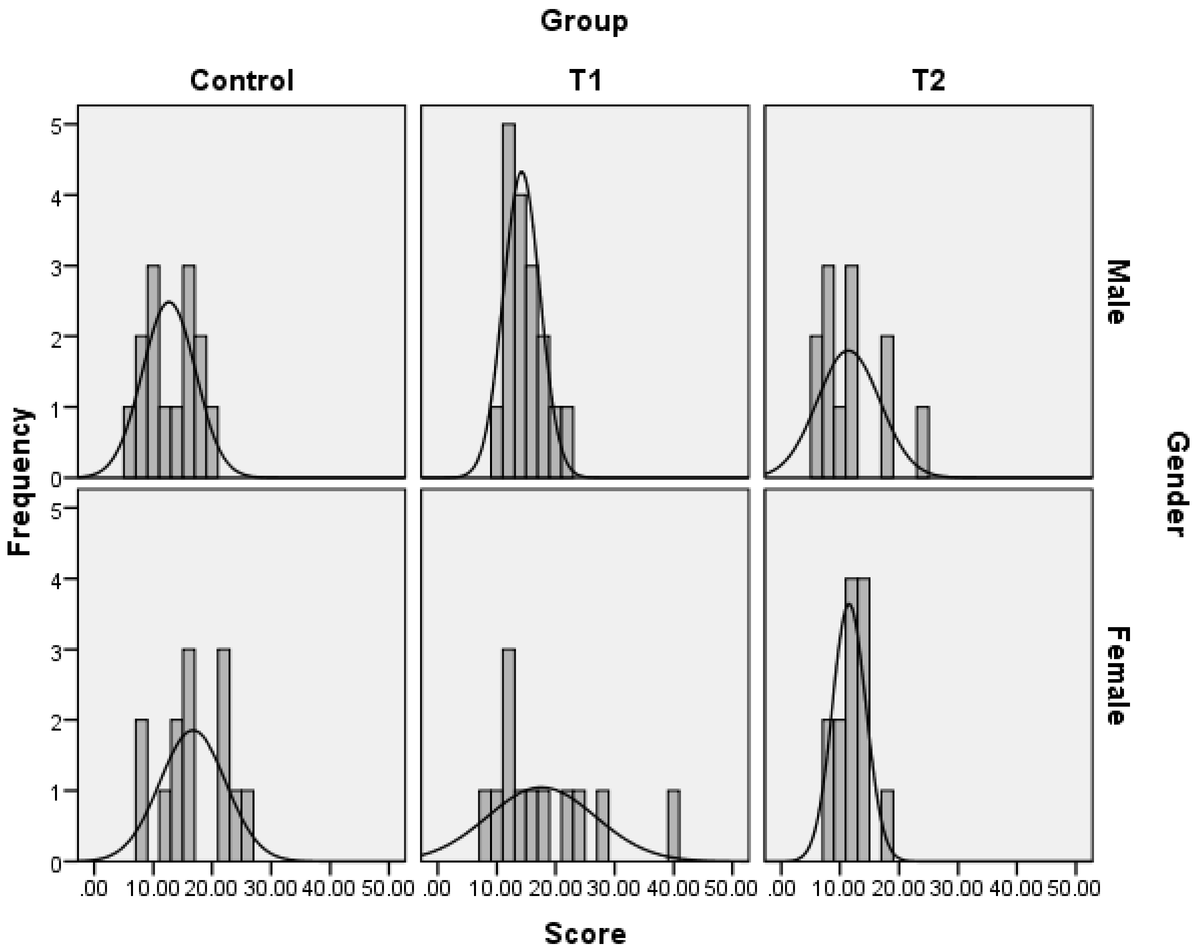


Figure 21. Histogram for Two-Way ANOVA Composite scores with classical adjustment only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Composite Only Two-Way ANOVA Including Gamer or Non-Gamer Results

A two-way between-groups analysis of variance was conducted to examine the effect of self-identifying as a gamer and music on achievement as measured by the composite of three criterion-referenced tests. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 35 displays full statistics for this test. The interaction effect between being a gamer and group was not statistically significant F (2, 75) =.67, p = .52. There was not a statistically significant main effect for being a gamer *F* (1, 75) = .03, *p* = .87 with no effect size (partial eta squared = .00). There was a statistically significant main effect for group *F* (2, 75) = 3.65, *p* = .03 with a small-medium effect size (partial eta squared = .09). The difference in mean scores was small (Table 36, Figure 22). Levene's Test (Table 37) was not significant (*p* = .483), indicating that the groups have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 38) indicated that the mean score for the Control (*M* = 14.59, *SD* = 5.37) was not significantly different from Treatment 1 (*M* = 15.74, *SD* = 6.65) or Treatment 2 (*M* = 11.67, *SD* = 4.04). LSD indicated that Treatment 1 was significantly different from Treatment 2 at $^{**}p < .01$. LSD indicated that Treatment 1 was significantly different from Treatment 2 at $^{**}p < .01$. The distribution curve for Figure 23 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 35

Test of Between-Subjects Effects for Two-Way ANOVA Composite Scores for Gamer or Non-Gamer Only

Dependent Variable: Scores on Composite Test							
Source	*Type III SS*	*df*	*MS*	*F*	*p*	*np*^{2}	*Observed Power* [b]
Corrected Model	279.835 [a]	5	55.967	1.838	.116	.109	.597
Intercept	14970.843	1	14970.843	491.564	.000	.868	1.000
Gamer	.822	1	.822	.027	.870	.000	.053
Group	222.009	2	111.005	3.645	.031	.089	.656
Gamer * Group	40.626	2	20.313	.667	.516	.017	.158
Error	2284.165	75	30.456				
Total	18440.000	81					
Corrected Total	2564.000	80					

[a] R^2 = .109 (Adjusted R^2 = .050)
[b] Computed using alpha = .05

Table 36

Descriptive Statistics for Two-Way ANOVA Composite Scores for Gamer or Non-Gamer Only

Dependent Variable: Scores on Composite Test

Group	Gamer or Non-Gamer	$\bar{x}^{a}$	σ	*n*
Control	Gamer	13.27	5.50	11
	Non-Gamer	15.50	5.27	16
	Total	14.59	5.37	27
T1	Gamer	16.11	4.91	9
	Non-Gamer	15.56	7.50	18
	Total	15.74	6.65	27
T2	Gamer	12.25	4.39	12
	Non-Gamer	11.20	3.82	15
	Total	11.67	4.04	27
Total	Gamer	13.69	5.04	32
	Non-Gamer	14.20	6.08	49
	Total	14.00	5.66	81

[a] 52 Items

Table 37

Levene's Test of Equality of Variances[a] *for Two-Way ANOVA Composite Scores for Gamer or Non-Gamer Only*

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
.905	5	75	.483

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Gamer + Group + Gamer * Group

Table 38

Post Hoc Tests / Multiple Comparisons for Two-Way ANOVA Composite Scores for Gamer or Non-Gamer Only

Dependent Variable: Scores on Composite Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	Control	T1	-1.1481	1.50199	.726	-4.7396	2.4433
		T2	2.9259	1.50199	.132	-.6655	6.5173
	T1	Control	1.1481	1.50199	.726	-2.4433	4.7396
		T2	4.0741*	1.50199	.022	.4827	7.6655
	T2	Control	-2.9259	1.50199	.132	-6.5173	.6655
		T1	-4.0741*	1.50199	.022	-7.6655	-.4827
LSD	Control	T1	-1.1481	1.50199	.447	-4.1403	1.8440
		T2	2.9259	1.50199	.055	-.0662	5.9180
	T1	Control	1.1481	1.50199	.447	-1.8440	4.1403
		T2	4.0741**	1.50199	.008	1.0820	7.0662
	T2	Control	-2.9259	1.50199	.055	-5.9180	.0662
		T1	-4.0741**	1.50199	.008	-7.0662	-1.0820

Based on observed means.
The error term is *MSE* = 30.456
* The mean difference is significant at the .05 level.
** The mean difference is significant at the .01 level.

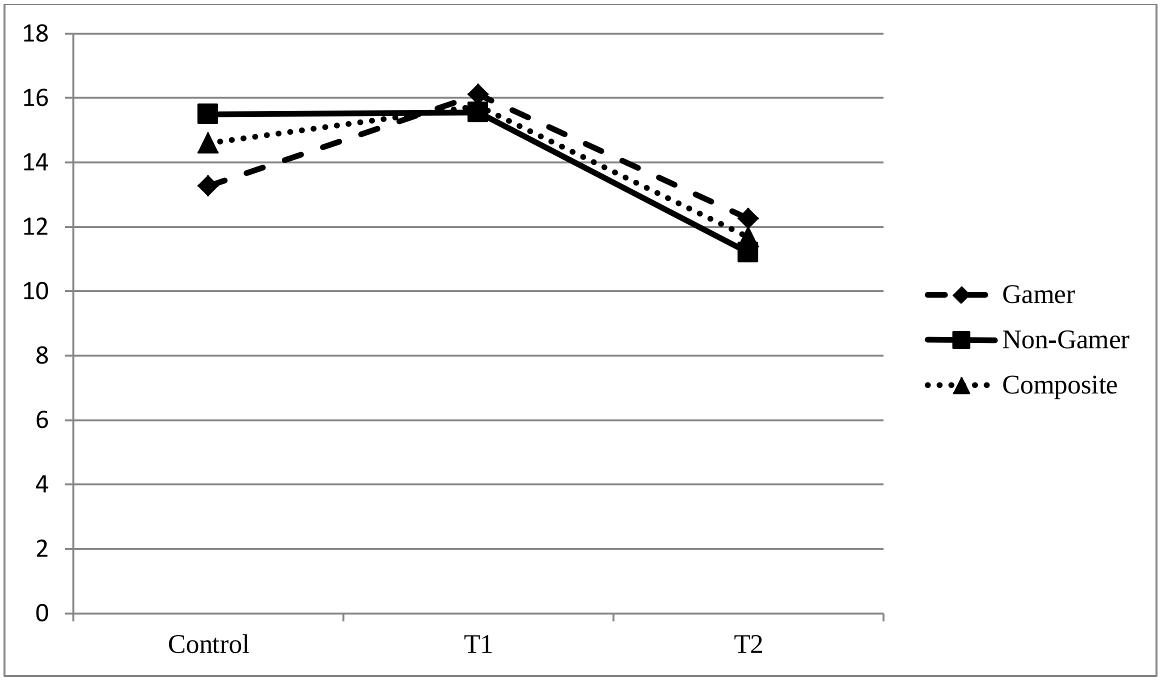


Figure 22. Means comparison for Two-Way ANOVA Composite scores including Gamer or Non-Gamer. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

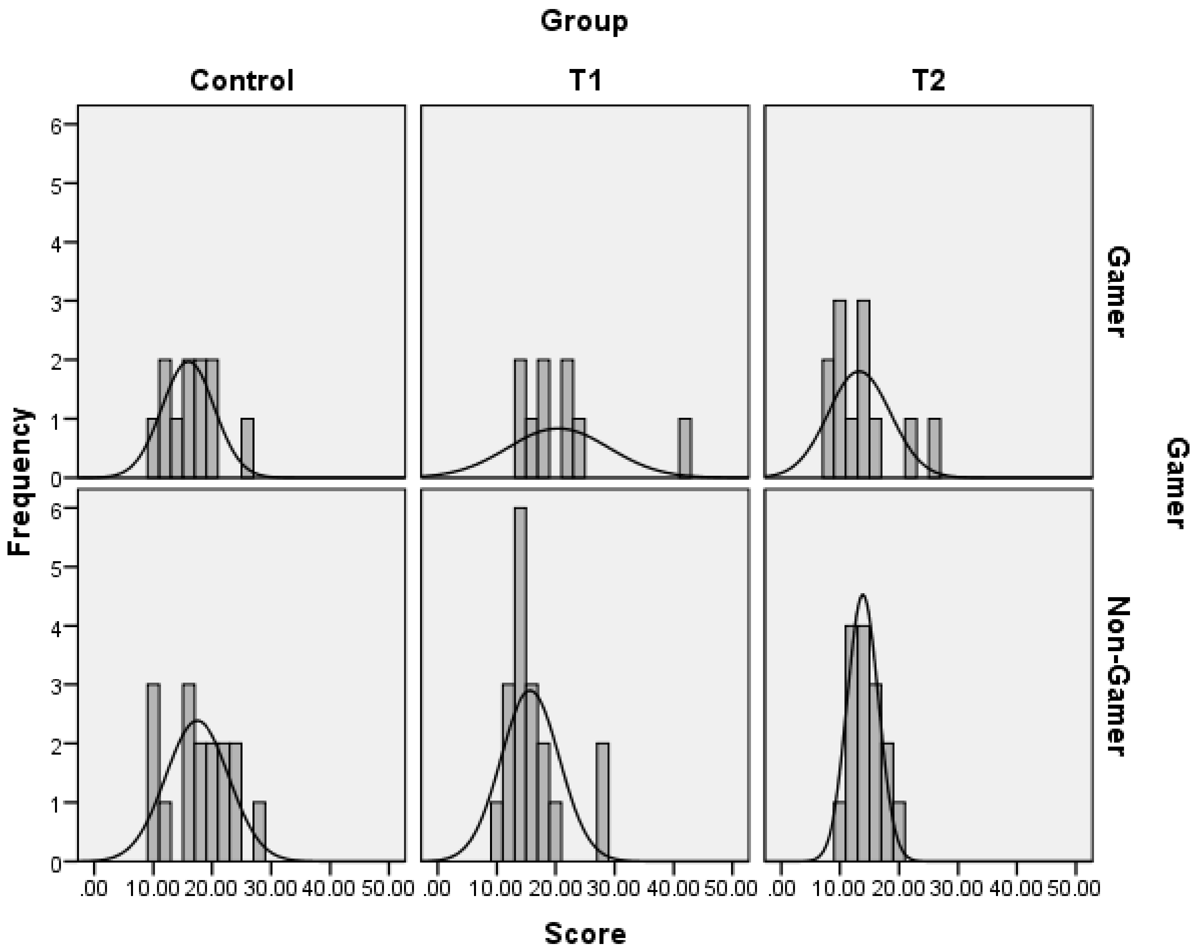


Figure 23. Histogram for Two-Way ANOVA Composite scores including Gamer or Non-Gamer. This Figure illustrates the distribution for scores in comparison to a normal curve.

Composite Only Three-Way ANOVA Including Gamer or Non-Gamer and Gender Results

A three-way between-groups analysis of variance was conducted to examine the effect of self-identifying as a gamer, gender, and music on achievement as measured by the composite of three criterion-referenced tests. Participants were randomly assigned to one of three groups: Control (no music), Treatment 1 (Mozart's *Sonata for Two Pianos in D Major*), or Treatment 2 (self-selected music). Table 39 displays full statistics for this test. The interaction effect between gender and group was not statistically significant F (2, 69) = 1.33, p = .27. The interaction effect between gender and being a gamer was not statistically significant F (1, 69) = .97, p = .33. The interaction effect between group and being a gamer was not statistically significant F (2, 69) = .14, p = .87. The interaction effect between gender, group, and being a gamer was not statistically significant F (2, 69) = 1.35, p = 2.66. There was not a statistically significant main effect for being a gamer F (1, 69) = 1.49, p = .23 with a small effect size (partial eta squared = .02). There was a statistically significant main effect for group F (2, 69) = 3.43, p = .04 with a small-medium effect size (partial eta squared = .09). There was a statistically significant main effect for gender F (1, 69) = 5.06, p = .03 with a small-medium effect size (partial eta squared = .07). The difference in mean scores was small (Table 40, Figures 24 and 25). Levene's Test (Table 41) was significant (p = .012) which indicates that the groups do not have equal variance; data were then transformed logarithmically in SPSS with the Lg10 transformation in order to better assess Levene's Test. After data transformation, Levene's Test (Table 42) was not significant (p = .13), indicating that the groups have equal variance. Post-hoc comparisons using the Tukey HSD and LSD tests (Table 43) indicated that the mean score for the Control (M = 14.59, SD = 5.37) was not significantly different from Treatment 1 (M = 15.74, SD = 6.65) or Treatment 2 (M = 11.67, SD = 4.04). LSD indicated that Treatment 1 was significantly different from Treatment 2 at **p < .01. The distribution curve for Figure 26 indicates that the data are not normally distributed; this is considered normal for social sciences research.

Table 44 shows how members of each gender, group, and gamer or non-gamer's observed mean scores on the composite test compare to observed means for Control, T1, and T2 without being broken down. Due to the small n for each of these groups, the data should be used for informational purposes only. Male gamers in T1 (M = 15.86) achieved higher scores on the composite test than the general population of all T1 members (M = 15.74). Male non-gamers in

T2 (*M* = 12.00) achieved higher scores on the composite test than the general population for all T2 members (*M* = 11.67). Females in both the Control and T1—regardless of being a gamer or non-gamer—scored higher on the composite test than the general population for these groups. The only place where females scored lower was in T2—specifically the non-gamer group (*M* = 10.80)—as compared to the general population of all T2 members (*M* = 11.67).

Table 39

Test of Between-Subjects Effects for Three-Way ANOVA Composite Only Gender, Group, and Gamer or Non-Gamer

Dependent Variable: Scores on Composite Test

Source	*Type III SS*	*df*	*MS*	*F*	*p*	*np*^{2}	*Observed Power* [b]
Corrected Model	545.301[a]	11	49.573	1.694	.093	.213	.784
Intercept	10395.155	1	10395.155	355.311	.000	.837	1.000
Gender	148.050	1	148.050	5.060	.028	.068	.602
Group	200.831	2	100.415	3.432	.038	.090	.626
Gamer	43.540	1	43.540	1.488	.227	.021	.225
Gender * Group	77.714	2	38.857	1.328	.272	.037	.278
Gender * Gamer	28.507	1	28.507	.974	.327	.014	.164
Group * Gamer	7.961	2	3.980	.136	.873	.004	.070
Gender * Group * Gamer	79.094	2	39.547	1.352	2.66	.038	.282
Error	2018.699	69	29.257				
Total	18440.000	81					
Corrected Total	2564.000	80					

[a] R^2 = .213 (Adjusted R^2 = .087)
[b] Computed using alpha = .05

Table 40

Descriptive Statistics for Three-Way ANOVA Composite Only Gender, Group, and Gamer or Non-Gamer

Dependent Variable: Scores on Composite Test					
Gender	Group	Gamer or Non-Gamer	$\bar{x}^a$	σ	*n*
Male	Control	Gamer	12.20	4.42	10
		Non-Gamer	13.75	5.19	4
		Total	12.64	4.50	14
	T1	Gamer	15.86	3.93	7
		Non-Gamer	12.88	2.03	8
		Total	14.27	3.33	15
	T2	Gamer	11.67	4.58	9
		Non-Gamer	12.00	6.28	5
		Total	11.79	5.01	14
	Total	Gamer	13.00	4.54	26
		Non-Gamer	12.82	4.14	17
		Total	12.93	4.34	43
Female	Control	Gamer	24.00	.	1
		Non-Gamer	16.08	5.38	12
		Total	16.69	5.60	13
	T1	Gamer	17.00	9.90	2
		Non-Gamer	17.70	9.57	10
		Total	17.58	9.16	12
	T2	Gamer	14.00	4.00	3
		Non-Gamer	10.80	2.15	10
		Total	11.54	2.85	13
	Total	Gamer	16.67	6.41	6
		Non-Gamer	14.94	6.83	32
		Total	15.21	6.71	38
Total	Control	Gamer	13.27	5.50	11
		Non-Gamer	15.50	5.57	16
		Total	14.59	5.37	27
	T1	Gamer	16.11	4.91	9
		Non-Gamer	15.56	7.50	18
		Total	15.74	6.65	27
	T2	Gamer	12.25	4.39	12
		Non-Gamer	11.20	3.82	15
		Total	11.67	4.04	27
	Total	Gamer	13.69	5.04	32
		Non-Gamer	14.20	6.07	49
		Total	14.00	5.66	81

[a] 52 Items

Table 41

Levene's Test of Equality of Variances[a] for Three-Way ANOVA Composite Only Gender, Group, and Gamer or Non-Gamer

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
2.449	11	69	.012

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Gender + Group + Gamer + Gender * Group + Gender * Gamer + Group * Gamer + Gender * Group * Gamer

Table 42

Levene's Test of Equality of Variances[a] for Three-Way ANOVA Composite Only Gender, Group, and Gamer or Non-Gamer Post Transformation

Dependent Variable: Scores on Composite Test

F	*df1*	*df2*	*p*
1.561	11	69	.130

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
[a] Design: Intercept + Group + Gender + Group * Gender

Table 43

Post Hoc Tests / Multiple Comparisons for Three-Way ANOVA Composite Gender, Group, and Gamer or Non-Gamer Only

Dependent Variable: Scores on Composite Test

	Group (I)	Group (J)	*M* Difference (I-J)	*SE*	*p*	95% Confidence Interval Lower Bound	Upper Bound
Tukey HSD	Control	T1	-1.1481	1.47212	.717	-4.6743	2.3780
		T2	2.9259	1.47212	.123	-.6003	6.4521
	T1	Control	1.1481	1.47212	.717	-2.3780	4.6743
		T2	4.0741*	1.47212	.020	.5479	7.6003
	T2	Control	-2.9259	1.47212	.123	-6.4521	.6003
		T1	-4.0741*	1.47212	.020	-7.6003	-.5479
LSD	Control	T1	-1.1481	1.47212	.438	-4.0850	1.7887
		T2	2.9259	1.47212	.051	-.0109	5.8627
	T1	Control	1.1481	1.47212	.438	-1.7887	4.0850
		T2	4.0741*	1.47212	.007	1.1373	7.0109
	T2	Control	-2.9259	1.47212	.051	-5.8627	.0109
		T1	-4.0741*	1.47212	.007	-7.0109	-1.1373

Based on observed means.
The error term is *MSE* = 29.257
* The mean difference is significant at the .05 level.
** The mean difference is significant at the .01 level.

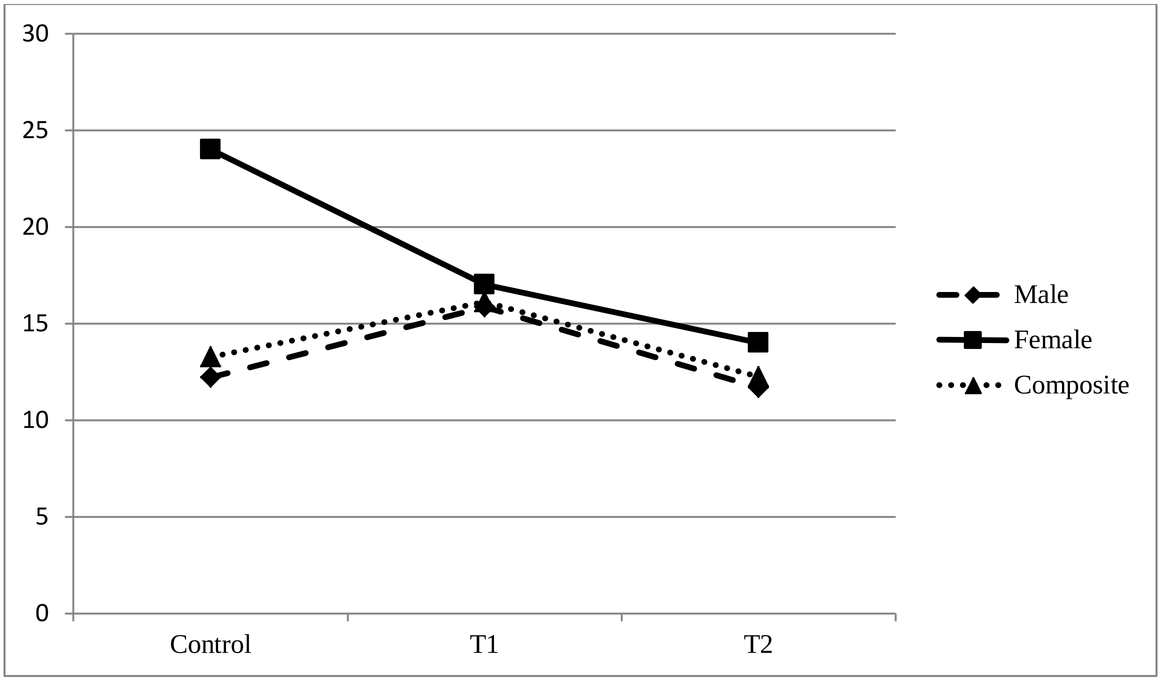


Figure 24. Means comparison for Three-Way ANOVA Composite scores for Gamers by gender only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

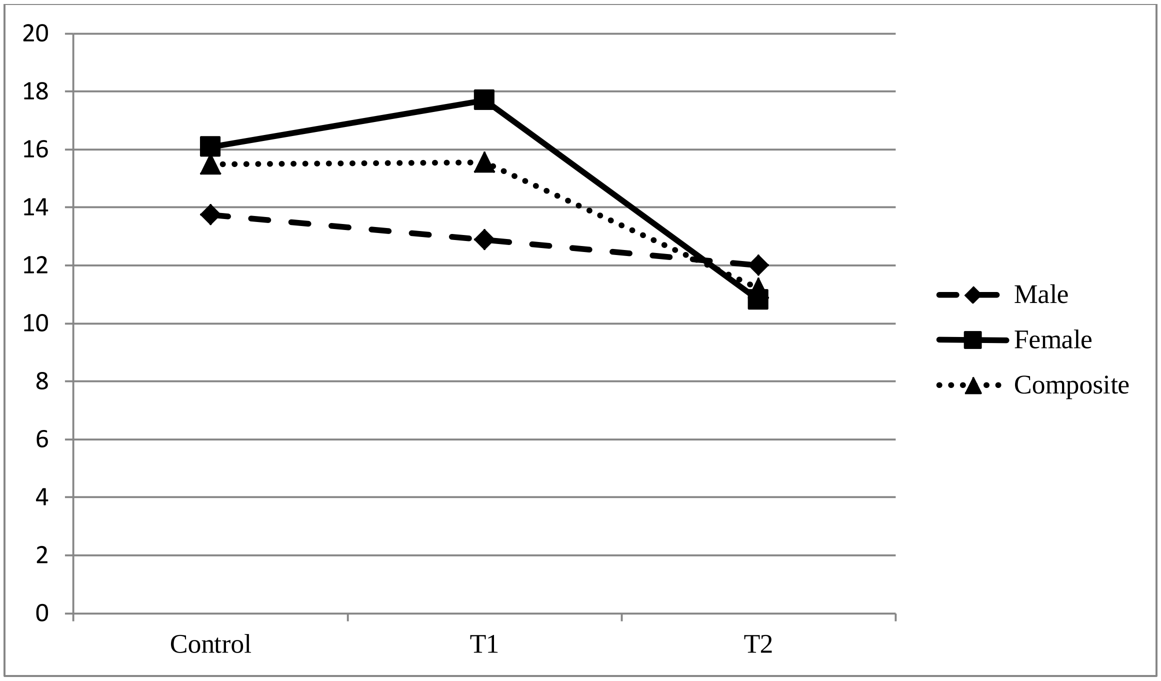


Figure 25. Means comparison for Three-Way ANOVA Composite scores for Non-Gamers by gender only. This Figure illustrates the distribution for male scores compared to female scores. An average score is also depicted.

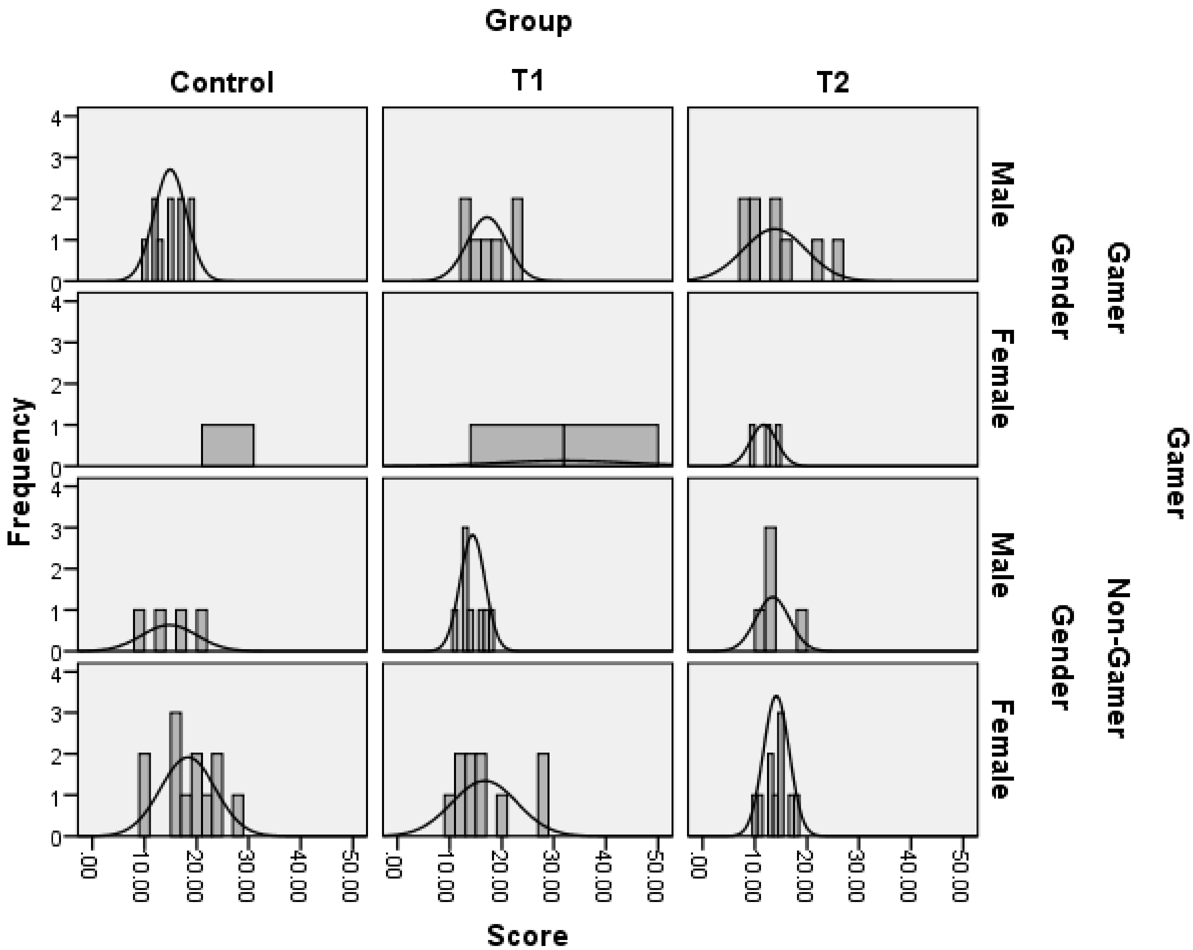


Figure 26. Histogram for Three-Way ANOVA Composite scores for Gamers and Non-Gamers by gender only. This Figure illustrates the distribution for scores in comparison to a normal curve.

Table 44

All Subjects Composite Scores by Test – Displayed by Gender, Group, and Gamer or Non-Gamer – As Compared to Overall Composite Score Mean

			Score[a]		Composite Score[a]	
Gender	Group	Gamer or Non-Gamer	*n*	$\bar{x}$	*n*	$\bar{x}$[b]
Male	Control	Gamer	10	12.20	27	14.59
		Non	4	13.75		
	T1	Gamer	7	15.86	27	15.74
		Non	8	12.88		
	T2	Gamer	9	11.67	27	11.67
		Non	5	12.00		
Female	Control	Gamer	1	24.00	27	14.59
		Non	12	16.08		
	T1	Gamer	2	17.00	27	15.74
		Non	10	17.70		
	T2	Gamer	3	14.00	27	11.67
		Non	10	10.80		
Total	Control	Gamer	11	13.27	27	14.59
		Non	16	15.56		
	T1	Gamer	9	16.11	27	15.74
		Non	18	15.56		
	T2	Gamer	12	12.25	27	11.67
		Non	15	11.20		

[a] 52 Items
[b] Mean composite for overall test.

Genres of Self-Selected Music for T2

Table 45 displays the genres of music that all participants selected for T2 (n = 27). This includes the group to which each subject was assigned (Control, T1, and T2), gender, number of subjects that selected a particular genre of music, and the observed means on the composite test for subjects that listened to that genre of music. 14 males and 13 females are represented. Due to the small number of participants relative to each genre of music, analysis must be done with caution. Composite scores by genre range from means of 8 to 15.5 correct answers with a median of 11. 6 genres (Rap / Hip-Hop, Alternative, R&B, Classic Rock, Crooner, and Dubstep) fell below the median with Classic Rock (*M* = 8, *SD* = N/A) scoring the lowest. Grindcore was on target with the median (*M* = 11, *SD* = N/A). 4 genres (Pop / Top 100, Punk, Classical, and Country) were above the median with Punk (*M* = 15.50, *SD* = 5.74) and Country (*M* = 15.50, *SD* = 3.54) tied with the highest scores, though Country had a smaller standard deviation. Figure 27 provides a graphic breakdown of genre by number of participants that selected each.

Table 45

Genres of Self-Selected Music Used for Treatment 2 With Mean Composite Scores

	Gender					*n*	Composite Score[a]	
Genre	Male	%	Female	%	Total %		$\bar{x}$	σ
Pop / Top 100	1	20	4	80	100	5	13.00	3.39
Rap / Hip-Hop	3	75	1	25	100	4	9.75	3.30
Punk	4	100	0	0	100	4	15.50	5.74
Alternative	2	50	2	50	100	4	8.50	2.38
Classical	2	100	0	0	100	2	14.00	1.41
Country	0	0	2	100	100	2	15.50	3.54
R&B	1	50	1	50	100	2	8.50	3.54
Classic Rock	0	0	1	100	100	1	8.00	N/A
Crooner	0	0	1	100	100	1	10.00	N/A
Dubstep	1	100	0	0	100	1	10.00	N/A
Grindcore	0	0	1	100	100	1	11.00	N/A

[a] 52 Items

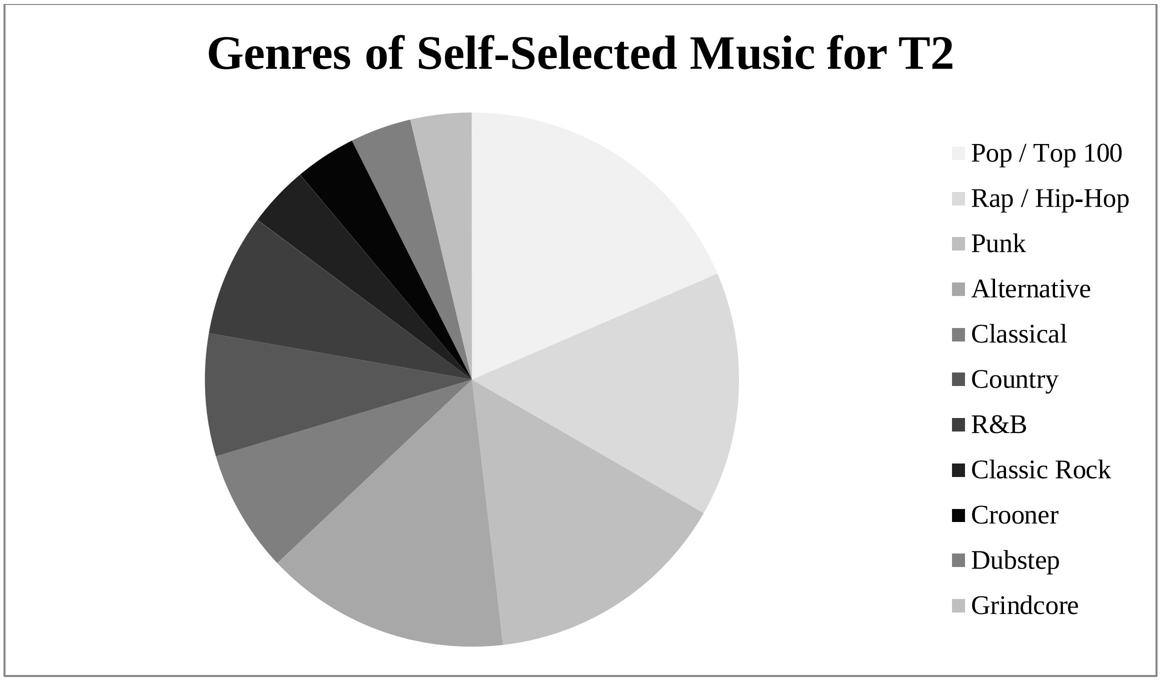


Figure 27. Pie Chart for genres of self-selected music that participants chose for use in T2. This Figure illustrates the relative size of each group's selection relative to other selections.

Subjects That Had More Than Three Classes About the Human Heart

Table 46 displays the group, major, mean scores on each test (identification, terminology, comprehension), and composite scores for all subjects that had more than three college classes that taught about the human heart (*n* = 3). It also includes mean scores for each test in the subjects own treatment group. All subjects were female. Two subjects (67%) were members of T1 and one subject (33%) was a member of T2. The instruction provided in this study should have served as a refresher for previously learned content for these three participants. The Nutrition major scored better on each test, and the composite, as compared to other members of T1. The Dance / Psychology major scored worse than other members of T1 on all but the terminology test. The Journalism major scored better than other members of T2 on all but the comprehension test.

Table 46

Subjects That Had >3 Classes That Taught About the Human Heart

Subject Group	Major	Identification Test[a]		Terminology Test[b]		Comprehension Test[b]		Composite Score[c]	
		Score	$\bar{x}^d$	Score	$\bar{x}^d$	Score	$\bar{x}^d$	Score	$\bar{x}^d$
Treatment 1	Nutrition	8.00	6.08	10.00	6.50	6.00	5.00	24.00	17.58
Treatment 1	Dance / Psychology	3.00	6.08	9.00	6.50	5.00	5.00	17.00	17.58
Treatment 2	Journalism	8.00	4.15	6.00	5.32	4.00	4.49	18.00	15.21

[a] 18 Items
[b] 17 Items
[c] 52 Items
[d] Mean of female subjects score on this test. Within own treatment groups only.
Note. All subjects that met criterion for inclusion in this table were female. No males met criterion.

Subjects That Self-Selected Classical Music during T2

Table 47 displays the mean scores on each test (identification, terminology, comprehension), and composite scores for all subjects that selected classical music for T2 (n = 2). It also includes mean scores for each test in T1 and T2. All subjects were male. These subjects are separated out because they selected a music genre that puts them in the same category as subjects in T1; as such, results may be expected to be similar. Both subjects scored higher on the composite (M = 13 and M = 14) than others in T2 (M = 11.79), though they were split compared to the mean of T1 (M = 14.27).

Table 47

Subjects That Self-Selected Classical Music During Treatment 2

Identification Test[a]			Terminology Test[b]			Comprehension Test[b]			Composite Score[c]		
Score	$\bar{x}^d$	$\bar{x}^e$	Score	$\bar{x}^d$	$\bar{x}^e$	Score	$\bar{x}^d$	$\bar{x}^e$	Score	$\bar{x}^d$	$\bar{x}^e$
4	5.20	3.57	3	4.33	3.29	6	4.73	4.93	13	14.27	11.79
4	5.20	3.57	4	4.33	3.29	7	4.73	4.93	15	14.27	11.79

[a] 18 Items
[b] 17 Items
[c] 52 Items
[d] Mean of male subjects score on this test within T1 only.
[e] Mean of male subjects score on this test within T2 only.
Note. All subjects that met criterion for inclusion in this table were male. No females met criterion.

Data Analysis Results

Research Question

Four null hypotheses will be tested in order to answer the following research question:

RQ: What are the effects of audio and gender within a 3D gaming environment on the achievement of different educational objectives?

Test of Null Hypothesis One

H_o1: There will be no significant difference in achievement on criterion tests between students that receive the default audio treatment and those that receive no audio treatment.

Using data analyzed earlier in this chapter, there was not a statistically significant difference in students' achievement on criterion tests depending on whether they received no audio treatment in the Control or the default audio treatment in T1 (Mozart's *Sonata for Two Pianos in D Major*). At a basic level of raw scores, Table 5 showed that the Control scores on the identification test (M = 5.70, SD = 2.83) differed from T1 identification scores (M = 5.59, SD = 3.30). Control scores on the terminology test (M = 4.44, SD = 2.76) differed from T1 terminology scores (*M* = 5.30, *SD* = 3.16). Control scores on the comprehension test (*M* = 4.44, *SD* = 2.45) differed from T1 comprehension scores (*M* = 4.85, *SD* = 2.93). And finally, composite test scores for the Control (*M* = 14.59, *SD* = 5.37) differed from T1 composite scores (*M* = 15.74, *SD* = 6.65).

Further analysis (Tables 16, 20, 25, 29, 34, 38, and 43) between the Control and T1 showed that although there was a difference in mean and standard deviation in both groups (with T1 generally scoring higher), the difference in scores between Control and T1 were not statistically significant. As such, H_o1 cannot be rejected. The hypothesis proper is not supported.

Test of Null Hypothesis Two

> **H_o2:** There will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive no audio treatment.

Using data analyzed earlier in this chapter, it is clear that there was a statistically significant difference in students' achievement on criterion tests depending on whether they received no audio treatment in the Control or the self-selected audio treatment in T2. At a basic level of raw scores, Table 5 showed that the Control scores on the identification test (*M* = 5.70, *SD* = 2.83) differed from T2 identification scores (*M* = 3.85, *SD* = 2.27). Control scores on the terminology test (*M* = 4.44, *SD* = 2.76) differed from T2 terminology scores (*M* = 3.64, *SD* = 1.66). Control scores on the comprehension test (*M* = 4.44, *SD* = 2.45) differed from T2 comprehension scores (*M* = 4.19, *SD* = 1.84). Finally, composite test scores for the Control (*M* = 14.59, *SD* = 5.37) differed from T2 composite scores (*M* = 11.67, *SD* = 4.04).

Further analysis (Table 16) shows that the difference between the Control composite scores and T2 composite scores was significant (*$p < .05$). Identification scores (Table 20) show that the difference between the Control and T2 scores was significant ($^*p < .02$). There was no significant difference between Control and T2 terminology scores (Table 25). There was no significant difference between Control and T2 comprehension scores (Table 29). When participants that selected classical music in T2 were grouped with T1, a significant difference was found (Table 34) between the Control and T2 ($^*p < .04$). Both Tables 38 ($p < .055$) and 43 ($p < .051$) found differences between the Control and T2 that would be significant at the ($p < .10$) level. Control and T2 showed a difference in mean and standard deviation in both groups (with T2 generally scoring lower). The difference in scores between Control and T2 were statistically significant for most tests—but most importantly on the composite score. As such, H_o2 can be rejected and the hypothesis proper is supported.

Test of Null Hypothesis Three

> **H_o3:** There will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive the default audio treatment.

Using data analyzed earlier in this chapter, it is clear that there was a statistically significant difference in students' achievement on criterion tests depending on whether they received the default audio treatment in T1 (Mozart's *Sonata for Two Pianos in D Major*) or the self-selected audio treatment in T2. At a basic level of raw scores, Table 5 showed that the T1 scores on the identification test (M = 5.59, SD = 3.30) differed from T2 identification scores (M = 3.85, SD = 2.27). T1 scores on the terminology test (M = 5.30, SD = 3.16) differed from T2 terminology scores (M = 3.64, SD = 1.66). T1 scores on the comprehension test (M = 4.85, SD = 2.93) differed from T2 comprehension scores (M = 4.19, SD = 1.84). And finally, composite test scores for T1 (M = 15.74, SD = 6.65) differed from T2 composite scores (M = 11.67, SD = 4.04).

Further analysis (Table 16) shows that the difference between T1 composite scores and T2 composite scores was significant (**$p < .01$). Identification scores (Table 20) show that the difference between T1and T2 scores was significant (*$p < .03$). Terminology scores (Table 25) show that the difference between T1 and T2 scores was significant (*$p < .02$). There was no significant difference between T1 and T2 comprehension scores (Table 29). When participants that selected classical music in T2 were grouped with T1, a significant difference was found (Table 34) between T1 and T2 (**$p < .01$). Tables 38 (**$p < .01$) and 43 (**$p < .01$) both found significant differences between the composite scores of T1and T2. T1and T2 showed a difference in mean and standard deviation in both groups (with T2 generally scoring lower). The difference in scores between T1 and T2 were statistically significant for most tests – but most importantly on the composite score. As such, H_o3 can be rejected and the hypothesis proper is supported.

Test of Null Hypothesis Four

H_o4: There will be no significant difference in achievement on criterion tests between males and females.

Using data analyzed earlier in this chapter, it is clear that there was a statistically significant difference in students' achievement on criterion tests depending on whether they were male or female. At a basic level of raw scores, Tables 6, 7, 13, 18, 22, and 27 showed that male (n = 43) scores on the identification test (M = 4.65, SD = 2.69) differed from female (n = 38) identification scores (M = 5.50, SD = 3.13). Male scores on the terminology test (M = 3.70, SD = 2.10) differed from female terminology scores (M = 5.32, SD = 3.00). Male scores on the comprehension test (M = 4.58, SD = 2.18) differed from female comprehension scores (M =

4.49, *SD* = 2.72). Finally, composite test scores for males (*M* = 12.93, *SD* = 4.34) differed from female composite scores (*M* = 15.21, *SD* = 6.71).

Further analysis (Table 12) shows that there was a statistically significant main effect for gender F (2, 75) = 3.97, *p = .05 with a small-medium effect size (partial eta squared = .05) for composite scores. Table 17 shows that there was not a statistically significant main effect for gender F (1, 75) = 1.91, p = .17 with a small effect size (partial eta squared = .03) for identification scores. Table 21 shows that there was a statistically significant main effect for gender F (1, 75) = 8.94, **p = .01 with a small-medium effect size (partial eta squared = .11) for terminology scores. Comprehension scores (Table 26) showed no statistically significant main effect for gender F (1, 75) =.10, p = .76 with no effect size (partial eta squared = .00).

Additional analysis adjusting for the two members of T2 that listened to classical music—by instead including them in T1—made no difference in the data (Table 31). Scores stayed the same. As before, there was a statistically significant main effect for gender F (1, 75) = 4.42, *p = .04 with a small effect size (partial eta squared = .06) as indicated in Table 30. Analysis looking at male gamers (n = 26, *M* = 13.00, *SD* = 4.54) and male non-gamers (n = 17, *M* = 12.82, *SD* = 4.14) composite scores when compared to female gamers (n = 6, *M* = 16.67, *SD* = 6.41) and female non-gamers (n = 32, *M* = 14.94, *SD* = 6.83) as displayed in Table 40 shows an interaction effect between gender and being a gamer was not statistically significant F (1, 69) = .97, p = .33 (Table 39). There was a statistically significant main effect for gender F (1, 69) = 5.06, *p = .03 with a small-medium effect size (partial eta squared = .07) also displayed in Table 39. Males and females showed a difference in mean and standard deviation in all groups (with females generally scoring higher). The difference in scores between males and females were statistically significant for most tests, but most importantly on the composite score. As such, H_o4 can be rejected and the hypothesis proper is supported.

Summary

This chapter described all data analysis conducted in order to test the four null hypotheses and broader research question stated in Chapter 1. Simple descriptive statistics combined with a one-way ANOVA, multiple two-way ANOVAs, and one three-way ANOVA were used to address each hypothesis. Tukey's HSD and LSD post hoc tests were used to find statistical significance. Results indicated that H_o1 was not rejected because the difference in achievement scores between Control and T1 were not statistically significant. H_o2, H_o3, and H_o4 were rejected, and thus the hypotheses supported, due to the statistical significance identified in the data. Therefore, there was a statistically significant difference in achievement scores between Control and T1 students, there was a statistically significant difference in achievement scores between T1 and T2 students, and there was a statistically significant difference in achievement scores between males and females. A more thorough examination of what this means is found in Chapter 5.

CHAPTER 5

CONCLUSIONS

Introduction

Digital educational games are expected to become a major part of K-12 and higher education in the near future. However, no framework exists for how best to design them, particularly in regard to music. This 2 X 3 factorial post-test only control experimental design study used two independent variables: audio (three levels: none, default, and self-selected) and gender (two levels: male and female) to measure how they affected the dependent variable of achievement scores on three criterion-referenced tests: identification, terminology, and comprehension. Participants played a short 3D digital educational game that addressed the human heart's functionality and parts based on content created by Dwyer and Lamberski (1977). The study examined what effects on achievement, if any, resulted from subjects listening to different types of music while playing the game and then being assessed on performance. This study also looked at whether or not there was a difference in this regard vis-à-vis gender. The overall results and conclusions of this study are presented in the rest of this chapter.

Summary of Findings and Discussion

This study of 81 college students—Control, T1, and T2 groups each with $n = 27$—at a mid-sized public university in Western Pennsylvania had interesting results that have presented as many questions as answers.

Null hypothesis one. This hypothesis—there will be no significant difference in achievement on criterion tests between students that receive the default audio treatment and those that receive no audio treatment—was not rejected. There was not a statistically significant difference in students' achievement on criterion tests between the Control (no audio) and T1

(Mozart's Sonata for Two Pianos in D Major) groups. This does not mean that no difference existed, however. On the identification test, the Control outperformed T1 by 1.97%. On the terminology test, the T1 group outperformed the Control by 19.40%. The comprehension test found the T1 group outperforming the Control by 9.23%. Finally, on the composite score for all tests, T1 outperformed the Control by 7.88% overall.

Null hypothesis two. This hypothesis—that there will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive no audio treatment—was rejected. There was a statistically significant difference in students' achievement on criterion tests between the Control (no audio) and T2 (self-selected music from Spotify) groups. The Control clearly scored more highly than the T2 group. On the identification test, the Control outperformed T2 by 48.05%. On the terminology test, the Control group outperformed T2 by 21.98%. The comprehension test found the Control once again outperforming T2, this time by 5.97%. Finally, on the composite score for all tests, the Control led T2 by 25.02% overall.

Null hypothesis three. This hypothesis—that there will be no significant difference in achievement on criterion tests between students that receive the self-selected audio treatment and those that receive the default audio treatment—was rejected. There was a statistically significant difference in students' achievement on criterion tests between the T1 (Mozart's Sonata for Two Pianos in D Major) and T2 (self-selected music from Spotify) groups. T1 clearly scored more highly than the T2 group. On the identification test, the T1 group outperformed T2 by 45.19%. On the terminology test, the T1 group outperformed T2 by 45.60%. The comprehension test found the T1 group once again outperforming T2, this time by 15.75%. Finally, on the composite score for all tests, T1 led T2 by 34.88% overall.

Null hypothesis four. This hypothesis—that there will be no significant difference in achievement on criterion tests between males and females—was rejected. There was a statistically significant difference in students' achievement on criterion tests depending on whether they were male (n = 43) or female (n = 38). Females clearly scored more highly than the male group. On the identification test, the female group outperformed the male group by 18.28%. On the terminology test, females outperformed males by 43.78%. The comprehension test found the males barely outperforming females, by 2%. Finally, on the composite score for all tests, females led males by 17.63% overall.

Research question. The RQ for this study—what are the effects of audio and gender within a 3D gaming environment on the achievement of different education objectives?—was answered. Overall, listening to Mozart's *Sonata for Two Pianos in D Major* improved scores on achievement tests slightly. However, it was not statistically significant. Subjects that were free to select their own music showed a declination in scores as compared to those that listened to silence or *Sonata for Two Pianos in D Major*. This was found to be statistically significant. *A posteriori*, the researcher decided to look at interaction effects between self-identifying as a gamer or non-gamer as it related to the function of being male or female and music type during treatment. Male gamers (n = 26) outperformed male non-gamers (n = 17) by 1.40% and female gamers (n = 6) outperformed female non-gamers (n = 32) by 11.58% on composite scores. Further, female gamers outperformed male gamers by 28.23% on composite scores.

Findings and Implications

Pragmatic and instructional findings. As established in Chapter 1, digital educational games are expensive to produce and music may be eliminated entirely if it is deemed to be unnecessary. The findings in this regard are mixed. Music affected performance. Mozart's music, in particular, resulted in improved scores; however, these were very modest improvements and not statistically significant. Participants that were given the opportunity to select their own music, on the other hand, showed statistically significant lower scores. The lower scores may be the result of decision fatigue and will be discussed in the Limitations section of this chapter. Overall, a clear answer regarding the efficacy of audio in digital educational game contexts was not found in this study.

In addition to the above practical findings, subjects that had received three or more classes that taught about the human heart prior to taking part in this study score more highly than those that had not had as much college level preparation. This may have been because the material in this test served as a review instead of initial contact with the information. Two members of T1 had more than three classes about the human heart prior to participation in this experiment, and one outperformed the others in her treatment group by 36.52%. The other member of T1 that fit this criterion barely underperformed compared to others in her group, by 3.41%. One member of T2 had more than three classes about the human heart prior to participation in the experiment. She outperformed the other members of T2 by 18.08%.

Finally, musical genre was suggested to have an impact on performance for members of T2. Based on the (admittedly limited) results of this study, punk and country music had a stronger positive influence on student achievement than other genres.

The researcher recommends that not giving students the opportunity to select their own music may be the correct choice when designing digital educational games. Although the best musical selections may well be different for various groups and depend on the content of the game, classical music appears to be a safe starting point for those interested in providing music. Though both punk and country music outperformed classical, the mean difference was not extreme and those two forms of music may be more divisive to learners. The approach utilized in this experiment may also be best as a review of material for an assessment as opposed to the initial method for presenting content.

Theoretical findings. As established in Chapter 1, information processing is the theoretical framework upon which this study is based. Some researchers suggest that there is no theoretical justification for including sound effects or music in educational media as it creates extraneous processing for the learner and may negatively affect performance (Clark & Mayer, 2008; Gredler, 2002). Clark and Mayer (2012) even found that participants that received no audio treatment performed 61%-149% more highly than those who received sound. The present study does not generally corroborate this assertion. While members of the Control (no audio) group did outperform the self-selected audio group (T2) by 25.02% overall, with a range of 5.97%-48.05% on individual tests, it was the T1 (Mozart) group that had the highest composite scores. T1 performed more highly on every test in comparison to the Control group – with a composite performance of 7.88% higher, and a range of 1.97%-19.40% on individual tests. While these differences were not found to be statistically significant, there is no doubt that those that received Mozart's *Sonata for Two Pianos in D Major* outperformed the silence group based on raw scores (as well as outperforming the group that self-selected their own music). Brown and Green's (2011) assertion that background music could have a positive effect on learner's information processing is thus validated.

Nantais and Schellenberg (1999) not only replicated Rauscher's Mozart Effect studies, but went on to conclude through further experimentation that as long as participants experienced their preferred listening condition, they would perform more highly than those that received their non-preferred listening condition. Given this, the findings in the present study were unexpected.

While participants in the present study did not necessarily take part in their preferred treatment group, those in T2 were able to self-select music, and they had lower performance on achievement tests than the Control and T1 groups.

Limitations

The study presented here has a number of limitations that need to be mentioned. Some of these were identified *a priori* though most were identified *a posteriori* after examining results and thinking critically about the design of the study. Some of these will be mentioned again in the section for recommendations for future research.

Generalizability. The participants in this study came from a convenience sample from a specific department at a mid-sized Western Pennsylvania university. As such, it is difficult to generalize the findings here to students at other universities or the general population.

Headphone selection. The headphones utilized in this study were disposable and inexpensive. This was due to the researcher having to pay out of pocket for them; affordability was an issue. Ideally, the headphones utilized in this study would have been of higher quality as audio fidelity may be a confounding variable in this experiment. Additionally, some participants expressed anger when asked to use the headphones supplied for the experiment; they wanted to use their own headphones and were upset that they could not. This frustration may have affected the results and is another possible confounding variable.

Music volume. Participants were asked to keep the volume level of the instructional materials between 40% and 80% of maximum volume (the default was 60%). Since different musical recordings have different perceived volume levels, this means that participants receiving music in T1 or T2 may have received louder or softer audio even in comparison to the volume level they set. This may have affected the participants' results.

Self-selected music choice. After completing the experiment, some students chose to speak with the researcher about the experiment and had questions about different aspects of it. One student from T2 mentioned that he was dissatisfied with his musical self-selection (Hip-Hop). As this was unexpected, the researcher asked why he listened to something he did not enjoy. His response was that he thought he would enjoy it, but Spotify did not play any songs that he liked and as a result he found the music frustrating. This was an unexpected possibility and yet another potential confounding variable, especially because this may have been true for other participants in T2.

Spotify as a self-selected music source. Participants in group T2 were limited to selecting music from Spotify as part of their treatment. Spotify has an extensive catalog of music to draw upon. However, some participants had never used it before and may have added to their cognitive load. Other participants may have preferred another music source (e.g. Pandora), which may have led to frustration on their part based on unfamiliarity with the platform or if their true preferences were not represented on Spotify. Though it would be difficult to account for this, it is still another factor that may have impacted results.

Self-selecting silence compared to self-selecting music. Fassbender, Richards, Bilgin, Thompson, and Heiden (2012) found that of their participants in a similar study to the one presented here, 65% preferred receiving instruction with music as opposed to 35% that preferred silence (p. 497). No data for this preference was recorded in the present study. Since preference for being assigned to a music or non-music group may have been a factor in performance, it is mentioned here as another potential limitation. Rauscher's own comments reinforce this idea, suggesting that results could be tied to whether or not the subject enjoys music (Spiegel, 2010).

Continuing music treatment throughout the assessment. The musical treatments in this study ended before participants started the achievement tests. The way in which people retrieve information is directly related to how it was encoded at the time it was processed (Clark & Mayer, 2008; Hannon & Craik, 2001; Tulving & Thomson, 1973). Therefore, it is possible that subjects in group T1 and T2 may have had difficulty in retrieving information since they were under different conditions from when the information was encoded. Note that the Control group received silence throughout both the treatment and the assessment.

Decision fatigue. There is a psychological phenomenon often referred to as decision fatigue. An excellent summary by Oto (2014) states that “our ability to force ourselves to do difficult things . . . draws upon a certain limited resource within us . . . when we’re forced to make tough decisions, it calls upon that same resource . . . [and] we start to make poor choices.” The connection to information processing and cognitive load (Kalyuga, 2009) then suggests that when participants were asked to self-select music as part of T2, they used up some of their decision-making resources. As such, this would also increase cognitive load and interfere with information processing. This could result in lower scores on achievement tests and is another confound.

Following directions. A few participants (n < 3) confided that they had not followed directions properly throughout the experiment. In one instance, a student completed the assessments before viewing the instructional materials. In the others, the students admitted to moving through everything as quickly as possible without reading or paying close attention – because they wanted to receive their food payment more quickly. While this could happen in any experiment, it is yet another confounding variable.

Repeated readings of instructional content. Participants were free to read through the instructional materials as many times as they felt necessary before moving on to the assessments. Based on the researcher's observations, most participants read through once. However, some did read the materials multiple times. This was not controlled for in the experiment and could be another area of concern.

Lack of animation or clear heart part labeling. The original "heart content" instructional materials (Dwyer and Lamberski, 1977) included images that clearly showed which section of the heart was being discussed at any one time. Other studies, such as one conducted by Ausman (2008) kept this idea intact via animation. Further, research by Daluz & Mapoy (2011) using similar methods found that subjects receiving both audio and video in unison scored more highly than those that did not receive them together. The present study eschewed these ideas in order to prevent graphics from being another variable in the experiment. It is the opinion of the researcher that this was a mistake as average scores on achievement tests in this experiment were much lower than those found in the work of Ausman (e.g. Ausman's participants in the Control scored an average of 51.77 points out of 60 [86.28%] as compared to the Control in the present study: 14.59 points out of 52 [28.06%]). In short, the instructional materials seem inappropriate *a posteriori*.

Lack of a pilot study. Discussions about whether or not a pilot was necessary took place early in the process of conceptualizing the present study. It was decided at that time that a pilot study was not necessary. It is the opinion of this researcher that this was a mistake as some of the other limitations addressed in this section would have been corrected had a pilot been conducted. Overall, not including this piece weakened the final product.

Gamers and non-gamers. The researcher chose *a posteriori* to look at how gender, group, and self-identifying as a gamer or non-gamer influenced performance. Overall, non-gamers (n = 49) outperformed gamers (n = 32) by 3.73% on composite scores. The number of participants in each sub-group has a lot of variance, however, and cannot adequately be compared against each other in a meaningful way (e.g. n = 26 for male gamers and n = 6 for female gamers). Nonetheless, this is something that could have influenced the outcomes. The researcher conjectures that the non-gamers may have found the instructional materials more novel than gamers and thus increased their attention.

Musicians compared to non-musicians. Previous work by Aheadi, Dixon, and Glover (2010) found that musicians did not respond to the Mozart Effect in the same way as non-musicians. Musicians performance remained stable regardless of exposure to music while non-musicians improved by more than 30% when listening to music. The present study did not collect data on whether or not the participants were musicians and this is another factor that could have influenced performance.

Recommendations for Future Research

Recommendations for future research were developed in part from the limitations of the present study as described earlier in this chapter. However, some items for future research were identified that came from both data analysis and discussions with colleagues that have studied similar research. Suggestions follow.

Increase sample size and use multiple universities. One limitation of the study is that a convenience sample of students from one particular university was utilized. Future research should involve a sample randomly drawn from multiple universities. This will increase generalizability of the data.

Higher quality sound. This same study should be repeated with higher quality headphones and lossless audio in order to examine whether or not audio fidelity plays a role in performance.

Control for music volume. Just as music fidelity may play a role in participant performance, so too may music volume cause an effect. This should be better controlled in future experiments.

Specific genres of music should be considered. Evidence suggests that certain aspects of music (e.g. pitch) may have an effect on learners (Ilie & Thompson, 2011). The present study suggests that different genres of music may have different effects on performance as well. As such, future research on specific musical genres should be conducted.

Music should be utilized throughout both the treatment and assessment. The Control group in this study received silence as part of the treatment and while participants filled out their respective assessments. The T1 and T2 groups did not listen to music while completing their respective assessments. As identified in the limitations section of this chapter, that means that information may have been encoded in a way that would be more difficult to decode and retrieve (Clark & Mayer, 2008; Hannon & Craik, 2001; Tulving & Thomson, 1973). This study should be repeated and correct this part of the method.

Greater emphasis on graphical content. Further to the suggestion that different content types should be explored, this same experiment should be repeated while providing greater visual cues as to which parts of the heart are being discussed at any given moment.

Lessen decision fatigue. As stated earlier in this chapter, one possible explanation for participants performing less well in T2 may be that they had to select a musical treatment from Spotify before continuing with the experiment. This may have increased cognitive load and reduced their ability to perform at a high level. As such, future experiments should allow participants to self-select the music of their choice a few days in advance of receiving the treatment. The appropriate music can then be queued up and waiting for them when they take part in the experiment, just as it was for participants in T1.

Allow participants to self-select treatment group. Participants were randomly assigned to the Control, T1, and T2 groups. However, research suggests that participants that can choose silence or music instead of having that decision made for them may show improved performance and, if nothing else, they prefer being given the option (Fassbender, Richards, Bilgin, Thompson, & Heiden, 2012). As such, future research should be conducted in which participants are permitted to self-select not only music type, but also whether or not they receive the music treatment at all.

Flow states should be measured. This suggestion was developed through conversations with Dr. Blair of Little Bird Games (http://littlebirdgames.com/), as his research on achievement systems in games took flow states into great consideration (Blair, 2011). Flow theory states that "experience will be most positive when a person perceives that the environment contains high enough opportunities for action (or challenges), which are matched with the person's own capacities to act (or skills)" (Csikszentmihalyi & LeFevre, 1989, p. 816). An educational game that appropriately addresses flow states may allow the participants to perform more highly as there are a variety of benefits associated with positive flow states (Blair, 2011, p. 13). Future research in this area should consider flow states in relation to music and achievement.

Different content types. The present study focused on the parts and function of the human heart, but perhaps due to lack of appropriate graphical representation, even the best performing subjects in the study did not do particularly well on the assessments. However, other types of instructional content that do not require graphical presentation may have proven more effective in this experiment. Future research should explore this possibility.

Compare musicians to non-musicians. As stated earlier, musicians and non-musicians responded to musical stimulus differently (Aheadi, Dixon, & Glover, 2010). It would be beneficial to repeat this study and focus on musician performance in comparison to non-musician performance.

Compare gamers to non-gamers. It was decided *a posteriori* to examine self-identified gamer's performance in comparison to non-gamer's performance. A larger study focusing on these two populations would help to extend knowledge in this area.

Summary

After reviewing the relevant literature on how audio and gender affect performance in 3D educational gaming environments and conducting a study intended to examine this very subject, the researcher concludes that music and gender both do play a role in participants achievement on criterion-referenced tests as they relate to the instructional content. The differences between performance by males and females were significant. Moreover, although the Control group did not produce achievement test results with statistical significance in comparison to those in the T1 group, it was clear that self-selecting music type had a negative impact on participant performance in T2 (in comparison to the Control and T1).

References

Aheadi, A., Dixon, P., & Glover, S. (2010). A limiting feature of the Mozart effect: Listening enhances mental rotation abilities in non-musicians but not musicians. *Psychology of Music, 38*(1), 107-117. doi:10.1177/0305735609336057

Almeida, L. C. (2008). The effects of different learning strategies to facilitate achievement of different educational objectives. *TechTrends: Linking Research & Practice to Improve Learning, 52*(3), 32-36. doi:10.1007/s11528-008-0152-y

Almeida, L. C. (2012). The effect of an educational computer game for the achievement of factual and simple conceptual knowledge acquisition. *Education Research International, 2012*, 5 pages. Retrieved from http://www.hindawi.com/journals/edu/2012/961279/ doi:10.1155/2012/961279

Amory, A. (2007). Game object model version II: A theoretical framework for educational game development. *Educational Technology Research & Development, 55*(1), 51-77. doi:10.1007/s11423-006-9001-x

Anderson, D. R., & Levin, S. R. (1976). Young children's attention to "Sesame Street." *Child Development, 47*(3), 806-811. doi:10.1111/1467-8624.ep12241861

Annetta, L. A. (2008). Video games in education: Why they should be used and how they are being used. *Theory Into Practice, 47*(3), 229-239. doi:10.1080/00405840802153940

Annetta, L. A. (2010). The "I's" have it: A framework for serious educational game design. *Review of General Psychology, 14*(2), 105-112. doi:10.1037/a0018985

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York, NY: Academic Press.

Aukerman, M. J. (1996). We [Recorded by Descendents]. On *Everything sucks* [CD]. Fort Collins, CO: The Blasting Room.

Ausman, B. D. (2008). *Effects of prior knowledge and multimodal presentation on various levels of educational objectives* (Doctoral dissertation). Retrieved from https://etda.libraries.psu.edu/paper/8188/

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 11*(4). 417-423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2010). Primer: Working memory. *Current biology, 20*(4), R136-R140.

Baddeley, A .D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 48-79). New York, NY: Academic Press.

Blair, L. (2011). *The use of video game achievements to enhance player performance, self-efficacy, and motivation* (Unpublished doctoral dissertation). University of Central Florida, Orlando, FL.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction, 4*(1), 11-44.

Bowman, M., Frame, D. L., & Kennette, L. N. (2013). Enhancing teaching and learning: How cognitive research can help. *Journal on Excellence in College Teaching, 24*(3), 7-28.

Bishop, M. J., Amankwatia, T. B., & Cates, W. M. (2008). Sound's use in instructional software to enhance learning: A theory-to-practice content analysis. *Educational Technology Research and Development, 56*(4), 467-486. doi:10.1007/s11423-006-9032-3

Bishop, M. J., & Cates, W. M. (2001). Theoretical foundations for sound's use in multimedia instruction to enhance learning. *Educational Technology Research and Development, 49*(3), 5-22.

Brown, A., & Green, T. D. (2011). *The essentials of instructional design: Connecting fundamental principles with process and practice* (2nd ed.). New York, NY: Pearson.

Cameron, B., & Dwyer, F. (2005). The effect of online gaming, cognition and feedback type in facilitating delayed achievement of different learning objectives. *Journal of Interactive Learning Research, 16*(3), 243-258.

Carter, C. W. (2012). Instructional audio guidelines: Four design principles to consider every instructional audio design effort. *TechTrends: Linking Research & Practice to Improve Learning, 56*(6), 54-58.

Chabris, C. F. (1999). Prelude or requiem for the 'Mozart effect'? *Nature, 400*(6747), 826-827.

Champsas, I., Leftheris, I., Tsiatsos, T., Terzidou, T., & Mavridis, A. (2012). OpenGames: A framework for implementing 3D collaborative educational games in OpenSim. *Proceedings of the European Conference on Games Based Learning,* 82-92.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210.

Clark, R. C., & Mayer, R. E. (2008). *E-Learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (2nd ed.). San Francisco, CA: Pfeiffer.

Clark, R. C., & Mayer, R. E. (2012). Using rich media wisely. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (Third edition. pp. 309-320). New York, NY: Pearson.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101.

Corbit, M. (2005). Moving into Cyberspace: Game worlds for learning. *Knowledge Quest, 34*(1), 18-22.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Cresswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Los Angeles, CA: Sage.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology, 56*(5), 815-822.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10-32. doi:10.1111/j.1467-8535.2009.01038.x

Daluz, C. T., & Mapoy, M. J. (2011). The effect of interactive media on elementary school childrens' story memory. *International Journal of Research & Review, 6*(1), 108-119.

de Freitas, S. I. (2006). Using games and simulations for supporting learning. *Learning, Media, & Technology, 31*(4), 343-358. doi:10.1080/17439880601021967

del Blanco, A., Torrente, J., Marchiori, E. J., Martinez-Ortiz, I., Moreno-Ger, P., & Fernández-Manjón, B. (2012). A framework for simplifying educator tasks related to the integration of games in the learning flow. *Educational Technology & Society, 15*(4), 305-318.

DiPietro, M., Ferdig, R. E., Boyer, J., & Black, E. W. (2007). Towards a framework for understanding electronic educational gaming. *Journal of Educational Multimedia and Hypermedia, 16*(3), 225-248.

Downs, E. P. (2008). How can Wii learn from video games? Examining relationships between technological affordances and socio-cognitive determinates on affective and behavioral outcomes (Doctoral dissertation). Retrieved from https://etda.libraries.psu.edu/paper/8893/

Driscoll, M. P. (2012). Psychological foundations of instructional design. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (Third edition. pp. 35-44). New York, NY: Pearson.

Dwyer, F. M. (1972). *A guide for improving visualized instruction*. State College, PA: Learning Services.

Dwyer, F. M., & Lamberski, R. (1977). *The human heart: Parts of the heart, circulation of blood and cycle of blood pressure*. State College, PA: Learning Services.

Egerfeldt-Nielsen, S. (2007). Third generation educational use of computer games. *Journal of Educational Multimedia and Hypermedia, 16*(3), 263-281.

Entertainment Software Assocation. (2013). *The Entertainment Software Association - Industry facts*. Retrieved from http://www.theesa.com/facts/

Ericsson, K., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245. doi:10.1037/0033-295X.102.2.211

Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43-71. doi:10.1002/piq.21143

Fassbender, E., Richards, D., Bilgin, A., Thompson, W. F., & Heiden, W. (2012). VirSchool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. *Computers & Education, 58*(1), 490-500. doi:10.1016/j.compedu.2011.09.002

French, D. P. (2006). Chapter 23: Animating your lecture. In J. J. Mintzes & W. H. Leonard (Eds.), *Handbook of College Science Teaching*. (pp. 223-231). Arlington, Virginia: National Science Teachers Association.

Gagné, R. M. (1970). Some new views of learning and instruction. *The Phi Delta Kappan, 51*(9), 468-472.

Gagné, R. M. (1974). Educational technology and the learning process. *Educational Researcher, 3*(1), 3-8.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction, 4*(1), 67-94.

Gelman, R. (1999). Cognitive development. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the cognitive sciences*. (pp. 128-130). Cambridge, Massachusetts: MIT Press.

Glassman, W. E., & Hadad, M. (2009). *Approaches to psychology* (5th ed.). London, England: Open University Press/McGraw-Hill Education.

Gredler, M. E. (2002). Games and simulations and their relationships to learning. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (Second edition. pp. 571-581). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gros, B. (2007). Digital games in education: The design of games-based learning environments. *Journal of Research on Technology in Education, 40*(1), 23-38.

Hallam, S., & Price, J. (1998). Can the use of background music improve the behavior and academic performance of children with emotional and behavioural difficulties? *British Journal of Special Education, 25*(2), 88-91.

Hallam, S., Price, J., & Katsarou, G. (2002). The effects of background music on primary school pupils' task performance. *Educational Studies, 28*(2), 111-122.

Hannon, B., & Craik, F. M. (2001). Encoding specificity revisited: The role of semantics. *Canadian Journal of Experimental Psychology*, *55*(3), 231-243. doi:10.1037/h0087369

Hoffman, L. (2009). Learning through games. *Communications of the ACM*, *52*(8), 21-22. doi:10.1145/1536616.1536624

Ilie, G., & Thompson, W. F. (2011). Experiential and cognitive changes following seven minutes exposure to music and speech. *Music Perception: An Interdisciplinary Journal*, *28*(3), 247-264.

Kalyuga, S. (2009). *Cognitive load factors in instructional design for advanced learners*. New York, NY: Nova Science Publishers.

Ke, F. (2008). Computer games application within alternative classroom goal structures: Cognitive, metacognitive, and affective evaluation. *Educational Technology Research & Development*, *56*(5/6), 539-556. doi:10.1007/s11423-008-9086-5

Keiser Family Foundation. (2010). *Generation M^2: Media in the lives of 8-to 18-year olds*. Retrieved from http://kaiserfamilyfoundation.files.wordpress.com/2013/04/8010.pdf

Killi, K., & Lainema, T. (2008). Foundation for measuring engagement in educational games. *Journal of Interactive Learning Research, 19*(3), 469-488.

Johnson, L., Adams, S., & Cummins, M. (2012) *NMC Horizon report: 2012 K-12 edition.* Austin, TX: The New Media Consortium.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013a) *NMC Horizon report: 2013 higher education edition*. Austin, TX: The New Media Consortium.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Ludgate, H. (2013b) *Technology outlook for community, technical, and junior colleges 2013-2018: An NMC Horizon Project sector analysis*. Austin, TX: The New Media Consortium.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014) *NMC Horizon report: 2014 K-12 edition*. Austin, TX: The New Media Consortium.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015) *NMC Horizon report: 2015 higher education edition*. Austin, TX: The New Media Consortium.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication, 50*(1), 46-70. doi:10.1111/j.1460-2466.2000.tb02833.x

Leonard, D. C. (2002). *Learning theories, A to Z*. Santa Barbara, CA: Greenwood Publishing Group.

Lin, H. (2006). The effect of questions and feedback used to complement static and animated visualization on tests measuring different educational objectives (Doctoral dissertation). Retrieved from https://etda.libraries.psu.edu/paper/7604/

Lok, C. (2005) The start of computer games. *Technology Review, 108*(6), 88.

Lockhart, R. S., & Craik, F. M. (1978). Levels of processing: A reply to Eysenck. *British Journal of Psychology, 69*(2), 171-175.

Mann, B. L. (2008). The evolution of multimedia sound. *Computers & Education, 50*(4), 1157-1173. doi:10.1016/j.compedu.2006.11.002

Martin, J., & Ewing, R. (2008). Power up! Using digital gaming techniques to enhance library instruction. *Internet Reference Services Quarterly, 13*(2/3), 209-225. doi:10.1080/10875300802103874

McAdams, S. & Bigand, E. (1993). Introduction to auditory cognition. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 1-9). New York. NY: Oxford University Press.

Merrill, M. D. (1983). Component display theory. In C. Reigeluth (Ed.), *Instructional-design theories and models* (pp. 279–333). Mahwah, NJ: Lawrence Erlbaum Associates.

Merrill, M. D. (1987). A lesson based on the Component Display Theory. In C. Reigeluth (Ed.),*Instructional theories in action: Lessons illustrating selected theories and models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97. doi:10.1037/h0043158

Moore, J. J. (2011). Behaviorism. *Psychological Record, 61*(3), 449-465.

Mustaquim, M. & Nyström, T. (2012). An inclusive framework for developing video games for learning. *Proceedings of the European Conference on Games Based Learning,* 348-355.

Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. *Psychological Science, 10*(4), 370-373.

O'Donohue, W., & Kitchener, R. F. (1999). Introduction: The behaviorisms. In R. F. Kitchener & W. T. O'Donohue (Eds.), *Handbook of Behaviorism*. (pp. 1-13). San Diego, CA: Academic Press.

Oto, B. When thinking is hard: Managing decision fatigue. *EMS World, 41*(5), 46-50.

Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5th ed.). New York, NY: McGraw-Hill.

Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*(3), 193-198. doi:10.1037/h0049234

Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning, 19*(1), 79-103. doi:10.1080/09588220600804087

Piccinini, G., & Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics, 37*(1), 1-38. doi:10.1007/s10867-010-9195-3

Potter, R. F., & Choi, J. (2006). The effects of auditory structural complexity on attitudes, attention, arousal, and memory. *Media Psychology, 8*(4), 395-419. doi:10.1207/s1532785xmep0804_4

Prensky, M. (2008). Students as designers and creators of educational computer games: Who else? *British Journal of Educational Technology, 39*(6), 1004-1019. doi:10.1111/j.1467-8535.2008.00823_2.x

Ragan, T. J. & Smith, P. L. (2004). Conditions theory and models for designing instruction. In D. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology* (pp. 623-649). Mahwah, NJ: Lawrence Erlbaum Associates.

Raghuvanshi, N., Lauterbach, C., Chandak, A., Manocha, D., & Lin, M. C. (2007). Real-time sound synthesis and propagation for games. *Communications of the ACM, 50*(7), 66-73.

Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation and Gaming, 23*(3), 261-276.

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature, 365*(6447), 611. doi:10.1038/365611a0

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1995). Listening to Mozart enhances spatial-temporal reasoning: Towards a neurophysiological basis. *Neuroscience Letters, 185*(1), 44-47. http://dx.doi.org/10.1016/0304-3940(94)11221-4

Reinard, J. C. (2006). *Communication research statistics*. Thousand Oaks, CA: Sage Publications.

Rey, G. D. (2011). Time advice and learning questions in computer simulations. *Australasian Journal of Educational Technology, 27*(3), 397-410.

Richards, D., Fassbender, E., Bilgin, A., & Thompson, W. F. (2008). An investigation of the role of background music in IVWs for learning. *ALT-J, Research in Learning Technology, 16*(3), 231-244.

Romney, P. (2003). Closing the achievement gap? *Independent School, 62*(4), 30.

Savan, A. (1999). The effect of background music on learning. *Psychology of Music, 27*(2), 138-146. doi:10.1177/0305735699272005

Schnotz, W., & Kürschner, C. (2007). A reconsideration of Cognitive Load Theory. *Educational Psychology Review, 19*(4), 469-508. doi:10.1007/s10648-007-9053-4

Schollmeyer, J. (2006). Games get serious. *Bulletin of Atomic Scientists, 62*(4), 34-39.

Schwartz, N. (2003). The impact of animation and sound effects on attention and memory processes. *Conference Papers -- International Communication Association,* 1-5.

Schwartz, N. (2005). Integral or irrelevant? The impact of animation and sound effects on attention and memory for multimedia messages (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Full Text database. (Order No. 3183508)

Sloman, S. (1999). Cognitive architecture. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the cognitive sciences*. (pp. 124-125). Cambridge, Massachusetts: MIT Press.

Spiegel, A. (2010, June 28). *Mozart effect, schmozart effect: Science misinterpreted : NPR*. Retrieved from http://www.npr.org/templates/story/story.php?storyId=128104580

Steele, K. M., Bella, S. D., Peretz, I., Dunlop, T., Dawe, L. A., Humphrey, G. K., Shannon, R. A., Kirby Jr., J. L., & Olmstead, C. G. (1999). Prelude or requiem for the 'Mozart effect'? *Nature, 400*(6747), 827.

Stevenson, J. W., & McCuistion, P. (1987). All-O-Gistics [Recorded by Descendents]. On *All* [CD]. Los Angeles, CA: SST.

St Clair-Thompson, H., Overton, T., & Botton, C. (2010). Information processing: A review of implications of Johnstone's model for science education. *Research in Science & Technological Education, 28*(2), 131-148. doi:10.1080/02635141003750479

Superannuation. (2014). *How much does it cost to make a big video game?* Retrieved from http://kotaku.com/how-much-does-it-cost-to-make-a-big-video-game-1501413649

The New Media Consortium. (n.d.). About the NMC | The New Media Consortium. Retrieved from http://www.nmc.org/about

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373. doi:10.1037/h0020071

U.S. Copyright Office. (2011). *U.S. Copyright Office – Copyright Law of the United States*. Retrieved from http://www.copyright.gov/title17/

Wang, L., Bastiaansen, M., Yang, Y., & Hagoort, P. (2011). The influence of information structure on the depth of semantic processing: How focus and pitch accent determine the size of the N400 effect. *Neuropsychologia, 49*(5), 813-820. doi:10.1016/j.neuropsychologia.2010.12.035

Watrin, J., & Darwich, R. (2012). On behaviorism in the cognitive revolution: Myth and reactions. *Review of General Psychology, 16*(3), 269-282. doi:10.1037/a0026766

Westera, W., Nadolski, R. J., Hummel, H. G. K., & Wopereis, I. G. J. H. (2008). Serious games for higher education: A framework for reducing design complexity. *Journal of Computer Assisted Learning, 24*(5), 420-432. doi:10.1111/j.1365-2729.2008.00279.x

Yu, D., & Yagle, J. (2013). *Using R for power analysis* [PDF document]. Retrieved from http://www.iup.edu/WorkArea/DownloadAsset.aspx?id=141665

Yu, X., Li, Z., Zhou, C., Zhuo, G., & Liu, Z. (2009). A design proposal of a game-based professional training system for highly dangerous professions. *Proceedings of the European Conference on Games Based Learning*, 388-394.

Yuxin, M., Williams, D., Prejean, L., & Richard, C. (2007). A research agenda for developing and implementing educational computer games. *British Journal of Educational Technology, 38*(3), 13-518. doi:10.1111/j.1467-8535.2007.00714.x

Zehnder, S. M., & Lipscomb, S. D. (2006). The role of music in video games. In P. Verderer & J. Bryant (Eds.). *Playing video games: Motives, responses, and consequences* (pp.241-258). Mahwah, NJ: Lawrence Erlbaum Associates.

Appendix A

Letter of Invitation

Title of Study: The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives.

Principal Investigator: Ryan L. Sittler

You have been invited to participate in a research study that will investigate the effects of audio and gender on the achievement of different educational objectives. Your participation in this study is voluntary. You can withdraw your consent and discontinue your participation in the project at any time by calling or emailing me (the Principal Investigator) using the contact information below. During the session you can terminate the experiment at any time by walking out of the lab. There will be no penalty or questions asked. No one will stop you from leaving and you will be free to leave at your discretion.

Once you arrive at the learning site, I will meet you and ask you to choose any computer in the lab. The computer will give you a brief description of what you are about to do. Once you are done reading the information, you will click ok for the experiment to begin. You will first complete a short demographic survey containing information about yourself. The entire experiment should last less than one hour.

In conducting this research I am very careful to safeguard your information. If you have any questions or concerns, contact either myself or my advisor:

Principal Investigator:

Principal Investigator
Ryan L. Sittler
Doctoral Candidate
Indiana University of Pennsylvania
Department of Communications Media
121 Stouffer Hall
1175 Maple Street
Indiana, PA 15705-1058
Phone: 610-780-2661
r.l.sittler@iup.edu

Faculty Sponsor/Committee Chair:

Dissertation Committee Chair
Mary Beth Leidman, Ed.D
Professor
Indiana University of Pennsylvania
Department of Communications Media
131 Stouffer Hall
1175 Maple Street
Indiana, PA 15705-1058
Phone: 724-357-5763
mbleid@iup.edu

If you are not interested in participating, simply hand your blank form back to me. If you aren't sure and have questions, please email me or talk to me in person – I'll be happy to answer any questions.

Appendix B

Volunteer Sign-Up Sheet

Student Name (Please Print)	
IUP Email Address	
Preferred Dates For Participation	

Appendix C

Voluntary Consent Form

Title of Study: The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives.

I have read and understand the information on the **Informed Consent Form** and **I consent to volunteer to be a subject in this study**. I understand that my personal information will be kept completely confidential and that I have the right to withdraw at any time. I certify that I have received a copy of the **Informed Consent Form** to keep in my possession. **I understand and agree to the conditions of this study as described**.

Name: (PLEASE PRINT) ____

Signature: ____

Date: ____

Phone Number: ____

IUP Email Address: ____

I certify that I have explained to the above individual the nature and purpose, potential benefits, and possible risks associated with participating in this research study. I have answered any questions that have been asked and have witnessed the above signature.

______________ _______________________________________

Date Primary Investigator's Signature

THIS PROJECT HAS BEEN APPROVED BY THE INDIANA UNIVERSITY OF PENNSYLVANIA INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS (PHONE 724.357.7730).

Appendix D

Informed Consent Form

Informed Consent Form for Social Science Research
Indiana University of Pennsylvania

Title of Project: The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives.

Principal Investigator: Ryan L. Sittler, Ph.D. Candidate
121 Stouffer Hall
Indiana, PA 15701
(610) 780-2661; r.l.sittler@iup.edu

1. **Purpose of the Study:** The proposed study will examine what role audio and gender play in the efficacy of a digital 3D game environment on the achievement of different educational objectives. This dissertation will extend pragmatic research on the use of sound in digital educational game design as well as to provide theoretical understanding of information processing for cognition and learning.

2. **Procedures to be followed:** This study will use a post-test experimental design (Cresswell, 2009) only. Subjects will be randomly assigned into three (3) groups. The control group will receive no audio during the instruction. Treatment group #1 will listen to Mozart's *Sonata for Two Pianos in D Major* during the instruction. Treatment group #2 will listen to the music of their choice during the instruction. This process will not take more than 1 hour. The study will compare differences between and among these three groups by utilizing test scores. Subjects will not write down their names or their student ID numbers. Therefore, there is no personal identifier associated with the data. The subjects will do the following on the day of the experiment. Subjects will have the procedure of the study explained by the principal investigator by reading the letter of invitation. Subjects will be asked to sign a Voluntary Consent Form. The principal investigator will have subjects sign two (2) copies of the informed consent form and subjects will keep one copy for themselves. Subjects will be asked to go to a computer lab according to the time slot they requested and will receive instruction on the contents of the human heart. Subjects will complete a brief demographic survey. After receiving the instruction, they will be asked to take three criterion-referenced tests.

3. **Discomforts and Risks:** There are no known psychological, personal, or academic risks associated with this study.

4. **Benefits:** The benefit for participation in this study will be that subjects will have a greater understanding of the human heart and 3D educational game environments. Societal benefits will stem from findings and add to the empirical research on this topic – specifically in regard to design and development of educational games and learning.

Informed Consent Form Continued

5. **Duration:** Subjects will be asked to spend one hour of their time to participate in this study.

6. **Statement of Confidentiality:** Subjects participation in this research study would remain confidential. The data will be stored and secured at a locked cabinet at the principal investigator's office on a Microsoft word protected file for three years. The files containing the data will be protected with a password. Only the principal investigator and his dissertation committee chair will have access to the records. The data will be destroyed in 2017. The following may review and copy records related to this research: The Office of Human Research Protections at Indiana University of Pennsylvania. Your confidentiality will be kept to the degree permitted by the technology used. No guarantees can be made regarding the interception of data sent via the Internet by any third parties.

7. **Right to Ask Questions:** Subjects can ask questions about this research. Contact the principal investigator at (610)-780-2661 if you have any questions about this study. Subjects can also call this number if he/she has concerns about this research, or if he/she feels that he/she has been harmed by this study. If subjects have questions about his/her rights as a research participant, or if subjects have concerns or general questions about the research, contact IUP's Office for Research Protections at (724)-357-7730. Subjects may also call this number if they cannot reach the principal investigator or wish to talk with someone.

8. **Payment for participation:** Subjects will receive food (e.g., pizza) for participating in the study.

9. **Voluntary Participation:** Subjects' decision to participate in this research is voluntary. Subjects can stop at any time. Subjects do not have to answer any questions they do not want to answer. Refusal to take part in or withdrawing from this study will involve no penalty or loss of benefits subjects would receive otherwise.

Principal Investigator:
Ryan L. Sittler
Doctoral Candidate
Indiana University of Pennsylvania
Department of Communications Media
121 Stouffer Hall
1175 Maple Street
Indiana, PA 15705-1058
Phone: 610-780-2661
r.l.sittler@iup.edu

Dissertation Committee Chair:
Mary Beth Leidman, Ed.D
Professor
Indiana University of Pennsylvania
Department of Communications Media
131 Stouffer Hall
1175 Maple Street
Indiana, PA 15705-1058
Phone: 724-357-5763
mbleid@iup.edu

If you agree to take part in this research study and the information outlined above, please sign your name and indicate the date below. Please keep a copy of the informed consent form for your records and return one signed copy of this form to the investigator.

____ ____________________

Person Granting Consent Date

If you are willing to participate in this study, please sign the statement below.

Informed Consent Form Continued

Voluntary Consent Form

Title of Study: The Effects of Audio and Gender Within a 3D Gaming Environment on the Achievement of Different Educational Objectives.

I have read and understand the information on the **Informed Consent Form** and **I consent to volunteer to be a subject in this study**. I understand that my personal information will be kept completely confidential and that I have the right to withdraw at any time. I certify that I have received a copy of the **Informed Consent Form** to keep in my possession. **I understand and agree to the conditions of this study as described**.

Name: (PLEASE PRINT) ____

Signature: ____

Date: ____

Phone Number: ____

IUP Email Address: ____

I certify that I have explained to the above individual the nature and purpose, potential benefits, and possible risks associated with participating in this research study. I have answered any questions that have been asked and have witnessed the above signature.

______________ ____
Date Primary Investigator's Signature

THIS PROJECT HAS BEEN APPROVED BY THE INDIANA UNIVERSITY OF PENNSYLVANIA INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS (PHONE 724.357.7730).

Appendix E

Demographic Survey

1. What is your age? _____________

2. What is your gender? Please check one. ☐ Male ☐ Female

3. What is your current GPA? ______________

4. What is your current year of college? Please check one.

☐ Freshman (Year One)
☐ Sophomore (Year Two)
☐ Junior (Year Three)
☐ Senior (Year Four)

5. Describe your academic history by selecting the option that best matches you. Please check one.

☐ I have been a Communications Media Major since I began taking classes at IUP.
☐ I changed my Major to Communications Media from another Major at IUP.
☐ I transferred to IUP from another school where I was a Communications Media Major.
☐ I transferred to IUP from another school where I was not a Communications Media Major.
☐ Other (Please Explain)

____

6. How many college courses have you taken that involved learning the workings of a human heart as part of the curriculum? Please check one.

☐ 0 – 2
☐ 3 – 5
☐ 6 or more

7. Do you own a cell phone? If yes, for how many years? (Leave blank if you do not own one). _______

8. Do you play games on your cell phone? Please check one. ☐ Yes. ☐ No.

9. Do you own a tablet device? If yes, for how many years? (Leave blank if you do not own one). _____

10. Do you play games on your tablet device? Please check one. ☐ Yes. ☐ No.

Demographic Survey Continued

11. **Would you describe yourself as a gamer? If yes, for how many years? (Leave blank for no). ______**

12. **How many hours per week do you play games (on any device)? _______**

13. **How many hours per day do you use your cell phone or tablet device for various functions (e.g., Twitter, Instagram, Snapchat, Facebook, Email, Web-browsing, Texting, Games, etc.) __________**

14. **If you listened to self-selected music, what did you listen to (e.g., band, genre)?_______**

15. **Do you play games on your tablet device? Please check one.** ☐ Yes. ☐ No.

16. **If this was not your first choice, what music would you have preferred to listen to (e.g., band, genre)?_______**

Appendix F

Instructional Script / Instruments

Note: The Instructional Script and three Criterion-Referenced Instruments (Identification, Terminology, and Comprehension tests) contained in this Appendix (pp. 145-162) are based on the work of Dwyer and Lamberski (1977) as adapted and modified by Almeida (2008). The materials used here are not identical to the originals but are quite close. Permission was sought from and granted by Dr. F. M. Dwyer to use/reproduce these materials within this dissertation. Evidence of permission can be found in Appendix G – Evidence of Permission to Reprint Materials.

Instructional Script

The Parts of the Heart

The human heart is a hollow, bluntly conical, muscular organ. Its pumping action provides the force that circulates the blood through the body. In the average adult, the heart is about five inches long and about two and one half inches thick. A man's heart weighs about 11 ozs. And a woman's heart weighs about 9 ozs.

The heart lies toward the front of the body and is in a slanting position between the lungs, immediately below the breastbone. The wide end points toward the right shoulder. The small end of the heart points downward to the front of the chest and towards the left. The lower portion of the heart is called the apex and is the part that you feel beating.

The human heart is really two pumps combined in a single organ which circulates blood to all parts of the body. The heart is divided longitudinally into two halves by the septum. The two halves may be compared to a block of two houses, which are independent of each other but have a common wall, the septum, between them. Each half of the heart is divided into an upper chamber and a lower chamber. The upper chambers on each side of the septum are called auricles; the lower chambers are called ventricles. Auricles have thin walls and act as receiving rooms for the blood, while the ventricles having thicker walls act as pumps moving the blood away from the heart. Although there is no direct communication between the right and left sides, both sides function simultaneously.

The heart contains several layers of membranes and muscle. The first set of membranes enclose the heart in a thin double-walled sac. The layer which forms the outer wall of the sac is called the pericardium. It is composed of a tough, transparent elastic tissue. It protects the heart from rubbing against the lungs and the walls of the chest. The inner portion of the double-walled sac is called the epicardium. It is attached to the heart muscle.

The heart muscle is called the myocardium; it controls the contraction and relaxation of the heart. The myocardium constitutes by far the greatest volume of the heart and its contraction is responsible for the propulsion of the blood through the body. The muscle varies in thickness; for example, the muscle in the auricle walls is thin when compared to the thickness of the muscle in the ventricle walls.

Finally, the endocardium is the name given to the membrane lining the inside of the heart wall. Blood enters the heart through veins. Only veins carry the blood to the heart. The superior and inferior vena cavas are the two veins which deposit blood in the right auricle; there are no valves at the opening of these veins.

The superior vena cava deposits blood into the right auricle from all body parts above heart level, for example, the head and arms. The inferior vena cava carries blood from parts of the body below heart level, for example, the trunk and legs, depositing the blood in the right auricle.

As blood from the body fills the right auricle, some of it begins to flow into the right ventricle immediately, through a common opening. This common opening, between the right auricle and the right ventricle, is called the tricuspid valve. This valve consists of three triangular flaps on thin, strong, fibrous tissue.

These flaps permit the flow of blood into the right ventricle, but prevent it from flowing backward into the right auricle because the ends of the flaps are anchored to the floor of the right ventricle by slender tendons.

Thus, blood passes from the right auricle through the tricuspid valve into the right ventricle. As the right ventricle is filled with blood, both ventricles begin to contract, creating pressure. While the blood pressure behind the tricuspid valve brings the flaps together and prevents the flow of blood between the right auricle and the right ventricle, the contraction of the right ventricle continues until the blood presses hard enough to open the pulmonary valve.

The pulmonary valve, located between the right ventricle and the pulmonary artery, consists of three flaps like the tricuspid valve. As soon as the right ventricle begins to relax from its contraction, the valve flaps are filled with blood backing up from the pulmonary artery. The flaps are pressed together, stopping the blood flow back into the right ventricle. The pulmonary valve only opens when the pressure in the right ventricle is greater than the pressure in the pulmonary artery, forcing the blood into the artery. In the pulmonary artery the blood is carried away from the heart to both the left and right lungs where it is cleaned and oxygenated.

Returning from the lungs, the blood enters the heart through four pulmonary veins and collects in the left auricle; these vein openings, like the vena cavas, have no valves. The left auricle then contracts when it is full, squeezing blood through the mitral valve into the left ventricle.

The mitral valve, located between the left auricle and the left ventricle, is smaller in construction to the tricuspid valve. As the left ventricle contracts simultaneously with its mate, the right ventricle, it forces blood behind the flaps of the valve thereby closing the passageway back to the left auricle. Like the tricuspid valve, the ends of the mitral valve flaps are anchored to the floor of the left ventricle by slender tendons.

The contraction of the left ventricle pumps the blood through the entire body. For this reason, it is the largest, strongest, and most muscular section of the heart. When the left ventricle is filled with blood, it contracts resulting in the pressure opening the aortic valve. The aortic valve is similar to the other flap like valves; the valve stops the backward flow of blood to the left ventricle and opens for the forward flow of blood to the aorta. The aorta is the large artery which carries the blood away from the heart back to the various parts of the body.

The Circulation of Blood Through the Heart

The directional flow of blood in the heart is determined by valves which allow the blood to flow in only one direction. These sets of valves are the tricuspid and mitral valves, which control the flow of blood from the auricles to the ventricles, and the pulmonary and aortic valves which control the flow of blood from the ventricles to the arteries. Both auricles receive blood simultaneously through vein openings which have no valves.

The right auricle receives its blood through the superior and inferior vena cavas, while the left auricle receives its blood through the pulmonary veins. A wave of muscular contractions starts at the top of the heart and passes downward, simultaneously, over both sides of the heart; that is, both auricles contract at the same time and then relax as the contraction passes down the ventricles. When the auricles are caused to contract, they become small and pale and in doing so the blood in their chambers is subjected to increased pressure which forces blood to the ventricles through the opened tricuspid and mitral valves.

As the ventricles fill, eddies of the blood float the flaps on both the tricuspid and mitral valves back to a partially closed position. The instant that the contraction of the auricles has been completed, the ventricles are stimulated to contract. This contraction increases the pressure in the ventricle chambers forcing the tricuspid and mitral valves completely closed, thereby preventing blood from being forced backwards into the auricles.

The auricles, relaxing from their contraction, receive a continuous blood flow from the vena cavas and the veins. As the ventricles continue to contract, pressure in these chambers force the pulmonary and aortic valves to open. The pulmonary valve, leading from the right ventricle, guards the entrance to the pulmonary artery. The aortic valve, leading from the left ventricle, guards the entrance to the aorta or aortic artery. Both are 3 flapped valves, and are together known as the semi-lunar valves. Prior to ventricle contraction, the valves are closed by back pressure provided by blood already in the exit arteries. When pressure in the ventricles becomes greater than that in the exit arteries due to ventricle contraction, the semi-lunar valves open.

With the semi-lunar valves open, blood flows from the right ventricle into the pulmonary artery on route to the lungs for cleaning and oxygen. Simultaneously, blood flows from the left

ventricle into the aorta for distribution throughout the entire body. Immediately following the pumping of blood into the arteries, the ventricles begin to relax. This relaxation lowers the pressure within their chambers and the greater pressure in the arteries close the semi-lunar valves. Pressure within the ventricles is sufficient, however, to maintain closure of the tricuspid and mitral valves against the already increasing auricle
pressure.

As the ventricles relax further, pressure within them rapidly decreases. At the same time, blood flowing into the auricles from the veins increases the auricle pressure. Due to the differential pressure between the auricles and ventricles, the tricuspid and mitral valves are forced partially open.

The circulation of blood through the heart begins again with the next auricle contraction. Auricle pressure fully opens the tricuspid and mitral valves resulting in a rapid flow of blood into the ventricles.

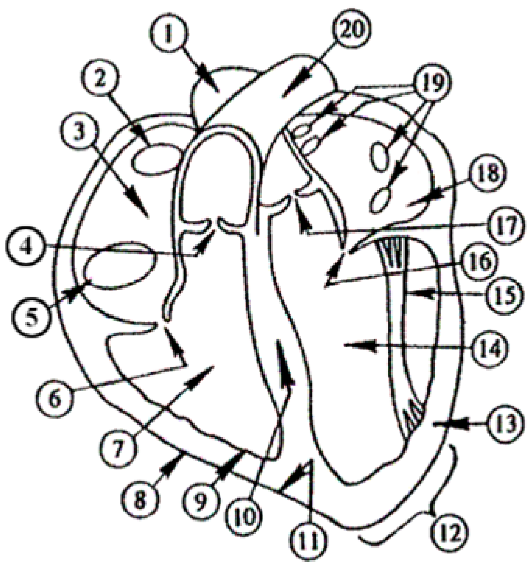
The Cycle of Blood Pressure in the Heart

The cycle of blood pressure in the heart consists of two distinct phases. One of these phases is called the diastolic or relaxation phase. In the diastolic phase, the heart relaxes between contractions. Blood flows into the heart, filling both auricles. While blood is flowing into the auricles, the arteries still maintain part of the pressure developed by a prior ventricle contraction. This is the time of lowest pressure in the arteries, or what is called the diastolic pressure.

During this phase the ventricles are also relaxing. The ventricles are slowly being filled with blood, due to the full auricles and partially opened tricuspid and mitral valves. The second phase, the systolic or contraction phase, begins when the auricles contract. The blood is forced through the tricuspid and mitral valves into the ventricles. The ventricles then contract forcing the blood through the semi-lunar valves into the pulmonary and aortic arteries.

The blood leaves the ventricles under terrific pressure and surges through the arteries with a force so great that it bulges their elastic walls. At this point, arterial blood pressure is greatest; we refer to this pressure as the systolic pressure. The heart begins to relax again. The semi-lunar valves are closed; blood flows into the auricles from the veins; and the tricuspid and mitral valves are forced partially open. The diastolic phase begins, and the cycle of blood pressure starts again.

Criterion-Referenced Identification Test Instrument



Arrow number one (1) points to the ___.

a. SEPTUM
b. AORTA
c. PULMONARY ARTERY
d. PULMONARY VEIN
e. NONE OF THESE

Arrow number two (2) points to the ___.

a. SUPERIOR VENA CAVA
b. INFERIOR VENA CAVA
c. PULMONARY ARTERY
d. TRICUSPID VALVE
e. AORTA

Arrow number three (3) points to the ___.

a. RIGHT VENTRICLE
b. RIGHT AURICLE
c. LEFT VENTRICLE
d. LEFT AURICLE
e. HEART MUSCLE

Arrow number four (4) points to the ___.

a. PULMONARY VALVE
b. PULMONARY VEIN
c. AORTA VALVE
d. TRICUSPID VALVE
e. MITRAL VALVE

Arrow number five (5) points to the ___.

a. AORTA
b. PULMONARY ARTERY
c. SUPERIOR VENA CAVA
d. INFERIOR VENA CAVA
e. PULMONARY VEIN

Arrow number six (6) points to the ___.

a. AORTA
b. PULMONARY VALVE
c. MITRAL VALVE
d. TRICUSPID VALVE
e. SEMI-LUNAR VALVE

Arrow number seven (7) points to the ___.

a. LEFT VENTRICLE
b. RIGHT VENTRICLE
c. RIGHT AURICLE
d. LEFT AURICLE
e. VASCULAR SPACE

Arrow number eight (8) points to the ___.

a. MYOCARDIUM
b. ECTODERM
c. PERICARDIUM
d. ENDOCARDIUM
e. EPICARDIUM

Arrow number nine (9) points to the ___.

a. ENDOCARDIUM
b. MYOCARDIUM
c. PERICARDIUM
d. ECTODERM
e. SEPTUM

Arrow number ten (10) points to the ___.

a. ENDOCARDIUM
b. PERICARDIUM
c. SEPTUM
d. MYOCARDIUM
e. AORTIC BASE

Arrow number eleven (11) points to the ___.

a. EPICARDIUM
b. PERICARDIUM
c. ENDOCARDIUM
d. MYOCARDIUM
e. NONE OF THESE

Arrow number twelve (12) points to the ___.

a. PERICARDIUM
b. MYOCARDIUM
c. ENDOCARDIUM
d. ENDODERM
e. APEX

Arrow number thirteen (13) points to the ___.

a. PERICARDIUM
b. ENDOCARDIUM
c. ECTOCARDIUM
d. ENDODERM
e. MYOCARDIUM

Arrow number fourteen (14) points to the ___.

a. RIGHT VENTRICLE
b. LEFT VENTRICLE
c. LEFT AURICLE
d. RIGHT AURICLE
e. APEX

Arrow number fifteen (15) points to the ___.

a. PULMONARY VEINS
b. TENDONS
c. AORTAS
d. PERICARDIUM
e. NONE OF THESE

Arrow number sixteen (16) points to the ___.

a. VENIC VALVE
b. PULMONARY VALVE
c. TRICUSPID VALVE
d. AORTIC VALVE
e. MITRAL VALVE

Arrow number seventeen (17) points to the ___.

a. SUPERIOR VENA CAVA
b. TRICUSPID VALVE
c. AORTIC VALVE
d. PULMONARY VALVE
e. MITRAL VALVE

Arrow number eighteen (18) points to the ___.

a. RIGHT AURICLE
b. RIGHT VENTRICLE
c. LEFT AURICLE
d. RIGHT AURICLE
e. SEMI-LUNAR CHAMBER

Arrow number nineteen (19) points to the ___.

a. INFERIOR VENA CAVA
b. SUPERIOR VENA CAVA
c. AORTAS
d. PULMONARY VEINS
e. PULMONARY ARTERIES

Arrow number twenty (20) points to the ___.

a. INFERIOR VENA CAVA
b. AORTA
c. PULMONARY ARTERY
d. SEPTUM
e. SUPERIOR VENA CAVA

Criterion-Referenced Terminology Test Instrument

___ is (are) the thickest walled chambers of the heart.

a. AURICLES
b. MYOCARDIUM
c. VENTRICLES
d. PERICARDIUM
e. ENDOCARDIUM

The contraction of the heart occurs during the ___ phase.

a. SYSTOLIC
b. SYMPATHETIC
c. DIASTOLIC
d. PARASYMPATHETIC
e. SYMPATRIC

Lowest blood pressure in the arteries occurs during the ___ phase.

a. SYMPATRIC
b. SYMPATHETIC
c. DIASTOLIC
d. SYSTOLIC
e. PARASYMPATHETIC

Blood from the right ventricle goes to the lungs through the ___.

a. TRICUSPID VALVE
b. AORTIC ARTERY
c. PULMONARY ARTERY
d. PULMONARY VEINS
e. SUPERIOR VENA CAVA

The ___ is (are) the strongest section(s) of the heart.

a. LEFT VENTRICLE
b. AORTA
c. SEPTUM
d. RIGHT VENTRICLE
e. TENDONS

When blood returns to the heart from the lungs, it enters the ___.

a. LEFT AURICLE
b. PULMONARY VALVE
c. LEFT VENTRICLE
d. RIGHT VENTRICLE
e. PULMONARY ARTERY

Vessels that allow the blood to flow from the heart are called the ___.

a. VEINS
b. ARTERIES
c. APEX
d. TENDONS
e. VALVES

Blood passes from the left ventricle out the aortic valve to the ___.

a. LUNGS
b. BODY
c. AORTA
d. PULMONARY ARTERY
e. LEFT AURICLE

The chamber of the heart which pumps oxygenated blood to all parts of the body is the ___.

a. RIGHT AURICLE
b. LEFT AURICLE
c. AORTA
d. LEFT VENTRICLE
e. RIGHT VENTRICLE

The ___ is another name for the part of the heart called the heart muscle.

a. APEX
b. EPICARDIUM
c. ENDOCARDIUM
d. MYOCARDIUM
e. SEPTUM

___ is (are) the part(s) of the heart which controls its contraction and relaxation.

a. MYOCARDIUM
b. ENDOCARDIUM
c. VENTRICLES
d. AURICLES
e. SEPTUM

The ___ is the name given to the inside lining of the heart wall.

a. EPICARDIUM
b. ENDOCARDIUM
c. PERICARDIUM
d. MYOCARDIUM
e. SEPTUM

Blood from the body enters the heart through the __.

a. AORTIC ARTERY
b. PULMONARY VEINS
c. PULMONARY ARTERY
d. SUPERIOR AND INFERIOR VENA CAVAS
e. SUPERIOR VENA CAVA ONLY

The membrane which borders on the inside lining of the pericardium and is connected to the heart muscle is called the ___.

a. EXTOXIM
b. EPICARDIUM
c. ENDOCARDIUM
d. MYOCARDIUM
e. ECTOCARDIUM

The ___ allow(s) blood to travel in one direction only.

a. SEPTUM
b. VALVES
c. ARTERIES
d. VEINS
e. TENDONS

The ___ is the common opening between the right auricle and the right ventricle.

a. MITRAL VALVE
b. TRICUSPID VALVE
c. SEPTIC VALVE
d. PULMONARY VALVE
e. AORTIC VALVE

The ___ is a triangular flapped valve between the left auricle and the left ventricle.

a. AORTIC VALVE
b. PULMONARY VALVE
c. SEPTIC VALVE
d. TRICUSPID VALVE
e. MITRAL VALVE

The semi-lunar valves are located at the entrance to the ___.

a. PULMONARY VEINS
b. SUPERIOR AND INFERIOR VENA CAVAS
c. PULMONARY AND AORTIC ARTERIES
d. MITRAL AND TRICUSPID VALVES
e. VENTRICLES

The outside covering of the heart is called the ___.

a. ENDOCARDIUM
b. EPICARDIUM
c. PERICARDIUM
d. MYOCARDIUM
e. NONE OF THESE

Immediately before entering the aorta, blood must pass through the ___.

a. LEFT VENTRICLE
b. MITRAL VALVE
c. LUNGS
d. SUPERIOR VENA CAVA
e. AORTIC VALVE

Criterion-Referenced Comprehension Test Instrument

Which valve is most like the tricuspid in function? ___.

a. PULMONARY
b. AORTIC
c. MITRAL
d. SUPERIOR VENA CAVA

When blood is being forced out the right ventricle, in which position is the tricuspid valve? ___.

a. BEGINNING TO OPEN
b. BEGINNING TO CLOSE
c. OPEN
d. CLOSED

When blood is being forced out the aorta, it is also being forced out of the ___.

a. PULMONARY VEINS
b. PULMONARY ARTERIES
c. SUPERIOR VENA CAVA
d. CARDIAC ARTERY

The contraction impulse in the heart starts in ___.

a. THE RIGHT AURICLE
b. BOTH VENTRICLES SIMULTANEOUSLY
c. BOTH AURICLES SIMULTANEOUSLY
d. THE ARTERIES

In the diastolic phase the ventricles are ___.

a. CONTRACTING, FULL OF BLOOD
b. CONTRACTING, PARTIALLY FULL OF BLOOD
c. RELAXING, FULL OF BLOOD
d. RELAXING, PARTIALLY FULL OF BLOOD

During the first contraction of the systolic phase, in what position will the mitral valve be? ___.

a. BEGINNING TO OPEN
b. OPEN
c. BEGINNING TO CLOSE
d. CLOSED

During the second contraction of the systolic phase, blood is being forced away from the heart through the ___.

a. PULMONARY AND AORTIC ARTERIES
b. SUPERIOR AND INFERIOR VENA CAVAS
c. TRICUSPID AND MITRAL VALVES
d. PULMONARY VEINS

When blood is entering through the vena cavas, it is also entering through the ___.

a. MITRAL VALVE
b. PULMONARY VEINS
c. PULMONARY ARTERY
d. AORTA

When the heart contracts, the ___.

a. AURICLES AND VENTRICLES CONTRACT SIMULTANEOUSLY
b. VENTRICLES CONTRACT FIRST, THEN THE AURICLES
c. RIGHT SIDE CONTRACTS FIRST, THEN THE LEFT SIDE
d. AURICLES CONTRACT FIRST, THEN THE VENTRICLES

While blood from the body is entering the superior vena cava, blood from the body is also entering through the ___.

a. PULMONARY VEINS
b. AORTA
c. INFERIOR VENA CAVA
d. PULMONARY ARTERY

When the blood leaves the heart through the pulmonary artery, it is also simultaneously leaving the heart through the ___.

a. TRICUSPID VALVE
b. PULMONARY VEINS
c. AORTA
d. PULMONARY VALVE

When the pressure in the right ventricle is superior to that in the pulmonary artery, in what position is the tricuspid valve? ___.

a. CLOSED
b. OPEN
c. BEGINNING TO CLOSE
d. CONFINED BY PRESSURE FROM THE RIGHT AURICLE

When the ventricles contract, blood is forced out the ___.

a. SUPERIOR AND INFERIOR VENA CAVAS
b. PULMONARY VEINS
c. TRICUSPID AND MITRAL VALVES
d. PULMONARY AND AORTIC VALVES

Blood leaving the heart through the aorta had left the heart previously through the ___.

a. VENA CAVAS
b. PULMONARY VEINS
c. PULMONARY ARTERY
d. TRICUSPID AND MITRAL VALVES

When the blood in the aorta is exerting a superior pressure on the aortic valve, what is the position of the mitral valve? **___**.

a. CLOSED
b. OPEN
c. BEGINNING TO OPEN
d. CONFINED BY PRESSURE FROM THE RIGHT VENTRICLE

When the tricuspid and mitral valves are forced shut, in what position is the pulmonary valve? **___**.

a. CLOSED
b. BEGINNING TO OPEN
c. OPEN
d. BEGINNING TO CLOSE

During the second contraction of the systolic phase, in what position is the aortic valve? **___**.

a. FULLY OPEN
b. PARTIALLY OPEN
c. PARTIALLY CLOSED
d. FULLY CLOSED

Blood is being forced out the auricles simultaneously as blood is **___**.

a. ENTERING ONLY THE VENA CAVAS
b. BEING FORCED OUT THE PULMONARY AND AORTIC VALVES
c. PASSING THROUGH THE TRICUSPID AND MITRAL VALVES
d. BEING FORCED OUT THROUGH THE PULMONARY ARTERY

If the aortic valve is completely open, the **___**.

a. SECOND CONTRACTION OF THE SYSTOLIC PHASE IS OCCURRING
b. DIASTOLIC PHASE IS OCCURRING
c. TRICUSPID AND MITRAL VALVES ARE COMPLETELY OPEN
d. BLOOD IS RUSHING INTO THE RIGHT AND LEFT VENTRICLES

When the heart relaxes, the **___**.

a. AURICLES RELAX FIRST, THEN THE VENTRICLES
b. RIGHT SIDE RELAXES FIRST, THEN THE LEFT SIDE
c. LEFT SIDE RELAXES FIRST, THEN THE RIGHT SIDE
d. VENTRICLES RELAX FIRST, THEN THE AURICLES

Appendix G

Evidence of Permission to Reprint Materials

Original material reprint permission email request from R. L. Sittler with private or non-essential information redacted:

From: Sittler, Ryan
To: fmd@psu.edu
Cc: zachary.stiegler@iup.edu
Subject: Requeting permission to reprint materials from: The Human Heart: Parts of the Heart, Circulation of Blood and Cycle of Blood Pressure.
Sent: Wed 3/4/2015 1:04 PM

Greetings, Dr. Dwyer.

[redacted] I'm a PhD Candidate at Indiana University of Pennsylvania in the Communications Media and Instructional Technology program and am preparing to defend my dissertation. I'm a former student of both Dr. Richard Lamberski and Dr. Luis Almeida. And it is my connection with them that has led me to you.

[redacted]

[redacted] One of my appendices contains the instructional script and assessment questions from the following document:

Dwyer, F.M., & Lamberski, R. (1977). *The human heart: Parts of the heart, circulation of blood and cycle of blood pressure.* State College, PA: Learning Services.

The Graduate School has instructed me to obtain permission from all copyright holders in order to include those materials in my dissertation. [redacted]

As such, do I have your persmission to reprint the instructional script and assessment questions from the above-mentioned document in my dissertation?

The requested permission extends to any future revisions and editions of my dissertation, including non-exclusive world rights in all languages, and to the prospective publication of my dissertation by UMI. These rights will in no way restrict republication of the material in any other form by you or by others authorized by you. Your permission will also confirm that you own [or your company owns] the copyright to the above-described material.

If these arrangements meet with your approval, please reply indicating so and include any special conditions that you may have.

[redacted]

If you have any questions or concerns, you are welcome to call me on my cell phone [redacted] or email me at this address. [redacted]

Thank you for your time, consideration, and attention.

Sincerely,
Ryan

Response material reprint permission granting email from Dr. F. M. Dwyer with private or non-essential information redacted:

From: FRANCIS M DWYER JR <fmd@psu.edu>
To: Sittler, Ryan
Cc:
Subject: Re: Requeting permission to reprint materials from: The Human Heart: Parts of the Heart, Circulation of Blood and Cycle of Blood Pressure.
Sent: Wed 3/4/2015 2:07 PM

Ryan:

I do own the copyright for the cited materials. You do have my permission to use the materials in the manner you suggest for your dissertation. If you plan further independent studies using these materials you would have to request permission at that time. I would suggest that you reference the materials appropriately since they are known widely [redacted]

[redacted] If at all possible when you have completed your dissertation I would appreciate having the opportunity to read the document.

Appendix H

Corrected Item-Total Correlations & Items Removed

Note: The composite test (pp. 164-166) is broken into 3 parts so that the table does not span multiple pages. Items identified for removal are highlighted in grey. Data for this analysis was used as a composite.

Item-Total Statistics

Question	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Ident 1	15.5802	29.622	.077	.	.636
Ident 2	15.5432	29.551	.085	.	.636
Ident 3	15.6173	29.039	.204	.	.628
Ident 4	15.6914	30.366	-.068	.	.644
Ident 5	15.6173	30.014	.002	.	.641
Ident 6	15.6420	29.808	.048	.	.638
Ident 7	15.6420	28.558	.319	.	.621
Ident 8	15.7284	29.700	.096	.	.635
Ident 9	15.5926	28.619	.284	.	.623
Ident 10	15.3951	28.642	.247	.	.625
Ident 11	15.6420	30.658	-.129	.	.648
Ident 12	15.3827	28.189	.334	.	.618
Ident 13	15.6914	29.366	.161	.	.631
Ident 14	15.6296	28.686	.285	.	.623
Ident 15	15.4198	29.322	.120	.	.634
Ident 16	15.6914	28.866	.278	.	.624
Ident 17	15.7037	28.736	.319	.	.622
Ident 18	15.6420	28.133	.414	.	.615
Ident 19	15.7284	29.275	.203	.	.629
Ident 20	15.7037	29.011	.252	.	.626

Item-Total Statistics

Question	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Termin 1	15.6173	30.364	-.069	.	.645
Termin 2	15.5802	28.697	.264	.	.624
Termin 3	15.5185	28.678	.252	.	.624
Termin 4	15.5802	29.422	.116	.	.634
Termin 5	15.7037	29.711	.085	.	.635
Termin 6	15.6790	29.121	.212	.	.628
Termin 7	15.4815	29.178	.150	.	.631
Termin 8	15.7778	30.500	-.107	.	.643
Termin 9	15.7037	29.411	.156	.	.631
Termin 10	15.6667	29.525	.115	.	.634
Termin 11	15.6420	29.108	.198	.	.629
Termin 12	15.5432	28.651	.262	.	.624
Termin 13	15.7037	29.336	.174	.	.630
Termin 14	15.6049	30.367	-.070	.	.645
Termin 15	15.4444	28.850	.209	.	.627
Termin 16	15.6049	29.192	.169	.	.630
Termin 17	15.6914	29.916	.034	.	.638
Termin 18	15.7160	29.031	.256	.	.626
Termin 19	15.7160	30.006	.018	.	.639
Termin 20	15.7037	29.011	.252	.	.626

Item-Total Statistics

Question	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Compreh 1	15.5185	29.928	.011	.	.641
Compreh 2	15.7037	29.661	.097	.	.635
Compreh 3	15.5802	28.897	.223	.	.627
Compreh 4	15.7037	29.361	.168	.	.631
Compreh 5	15.6543	29.904	.030	.	.639
Compreh 6	15.6420	29.133	.193	.	.629
Compreh 7	15.6049	28.942	.221	.	.627
Compreh 8	15.7531	29.988	.031	.	.638
Compreh 9	15.8025	29.385	.248	.	.628
Compreh 10	15.5679	30.148	-.028	.	.643
Compreh 11	15.6049	29.142	.179	.	.630
Compreh 12	15.4568	29.201	.144	.	.632
Compreh 13	15.6790	29.371	.155	.	.631
Compreh 14	15.6543	29.004	.226	.	.627
Compreh 15	15.6049	31.342	-.260	.	.657
Compreh 16	15.6296	29.361	.140	.	.632
Compreh 17	15.7037	30.586	-.119	.	.646
Compreh 18	15.5062	29.828	.030	.	.640
Compreh 19	15.6173	29.664	.074	.	.636
Compreh 20	15.4938	29.278	.132	.	.633