THE EFFECTS OF NON-SPEECH AUDIO IN THE INTERACTIVE ONLINE LEARNING ENVIRONMENT

A Dissertation

by

YUN LI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Oi-Man Kwok
Co-Chair of Committee,	Sherman Finch
Committee Members,	Jeffrey Liew
	Noelle Sweany
	Meta Rousseau
Head of Department,	Shanna Hagan-Burke

December 2018

Major Subject: Educational Psychology

ABSTRACT

Since the invention of the first recording device in the 1800s, the film and music industries began designing and implementing non-speech audio in their production process. In the field of education, sound was intensively used to facilitate learning since the expansion of computer technology in the 1970s. While the majority of research done and continues to focus on the use of speech audio, little attention has been paid to non-speech sounds. Therefore, this study endeavors to discover what non-speech audio can impact on learning process through effective design, developing, and implementing in an online learning environment.

A systematic literature review, presented in chapter 2, offers a thorough understanding of the circumstances in which non-speech audio is being used in the field of education. In addition, a model of the design of non-speech audio to support learning is also suggested based on experiential learning theory, learner-environment interactivity, and stimulus-response theory. After establishing the methodology of this study, the research questions and hypotheses are answered through the analysis of data collected during the experiment. Special emphasis is placed on learner motivation, online learning experience, and time spent in a custom-developed learning program.

The findings of this study reveal that the systematic application of non-speech audio will enhance the learning experience and stimulate learner motivation in online learning environments. Although unavoidable limitations are involved, the findings of

this study provide empirical evidence that may apply to the future design and development of non-speech audio in interactive online learning environments.

DEDICATION

To my dearest parents (Jianjun Li and Ronghua Yu) and myself

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Oi-Man Kwok, co-chair, Mr. Sherman Finch, and my committee members Dr. Jeffrey Liew, Dr. Noelle Sweany, and Dr. Meta Rousseau. Their support and guidance inspired me to overcome challenges and appreciate them as opportunities for learning. Thanks also go to Dr. Suzanna Ramos, Dr. Hector Ramos, and Dr. Joyce Juntune who enthusiastically helped me recruit the participants of this study.

I would also like to thank all of those who devoted their spare time to designing and developing the Learning Chinese program used in this study: Carolina Pereira, Cece George, Lauren Toler, Maggie Crosby, Reagan Woodlock, Ryan Canales, and Sarah Brown of the Department of Visualization at Texas A&M University, Tao Liu of the Department of Computer Science at Maharishi University of Management, and Junqi Yang of the Department of Computer Science at Texas A&M University.

Last but not least, thanks to my parents for their unconditional love, and my friends for their encouragement, support, and patience throughout the dissertation process.

CONTRIBUTORS AND FUNDING SOURCES

This dissertation was supported by a dissertation committee consisting of Professor Oi-man Kwok, Chair from the department of Educational Psychology; Professor Sherman Finch, Co-chair from the department of Visualization; Professor Noelle Sweany, Department of Educational Psychology; Professor Jeffrey Liew, Department of Educational Psychology; Professor Meta Rousseau, Department of Educational Psychology.

No outside funding was received for the research and compilation of this document.

NOMENCLATURE

API	Application Programming Interface
ASC	Academic Search Complete
ELT	Experiential Learning Theory
EPSY	Educational Psychology
INST	Instructional Strategy
LMS	Learning Management System
MANOVA	Multivariate Analysis of Variance
MIDI	Musical Instrument Digital Interface
SDT	Self-determination Theory
SIMS	The Situational Motivation Scale
S-R	Stimulus-Response
UEQ	User Experience Questionnaire

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

CONTRIBUTORS AND FUNDING SOURCES vi

NOMENCLATURE vii

TABLE OF CONTENTS viii

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER I INTRODUCTION 1

1.1 Statement of the Problem 2
1.2 Theoretical Framework 3
1.3 Purpose of the Study 4
1.4 Significance of the Study 5
1.5 Organization of the Study 5
1.6 Definition of Terms 6

CHAPTER II LITERATURE REVIEW 7

2.1 Theoretical Framework 7
2.2 Systematic Review Procedure 13
2.3 Results 15
2.4 Discussion 23

CHAPTER III METHODOLOGY 30

3.1 Experimental Research Design 30
3.2 Population and Participants 32
3.3 Instructional Program – Learning Chinese 33
3.4 Data Collection Procedure 44
3.5 Instrumentation 45

3.6 Data Analysis 50

CHAPTER IV RESULTS 54

4.1 A Report of Cronbach's Alpha 54
4.2 Descriptive Statistics...54
4.3 Checking the Assumptions of MANOVA 57
4.4 Summary 60

CHAPTER V CONCLUSION 61

5.1 A Summary of the Major Findings 62
5.2 Implications of the Study 66
5.3 Limitation of the Study 68
5.4 Recommendations for the Future Research 69

REFERENCES..71

APPENDIX A SYSTEMATIC REVIEW SEARCH TERMS 82

APPENDIX B EIGHT CHARACTERS IN THE PRACTICE SECTION......................84

APPENDIX C LEARNER CHARACTERISTICS QUESTIONS 91

APPENDIX D USER EXPERIENCE QUESTIONNAIR (UEQ) 92

APPENDIX E THE SITUATIONAL MOTIVATION SCALE (SIMS) 95

APPENDIX F SPSS OUTPUT OF MANOVA 96

APPENDIX G INVITATION EMAIL 100

APPENDIX H RESEARCH INSTRUCTION...102

LIST OF FIGURES

Page

Figure 1. Conceptual Model of Non-Speech Audio Design in the Interaction 10

Figure 2. Result of Systematic Review 16

Figure 3. One Learning Segment in Linear Learning Environment.................................27

Figure 4. Experimental Research Design Process 32

Figure 5. Instructional Content of Historical Overview...35

Figure 6. An Example of Interactive Writing Canvas in Practice Section.......................36

Figure 7. Matching Section 37

Figure 8. An Example of Stroke Order Assessment 38

Figure 9. An Example of the Practice Section 39

Figure 10. The Implementation of Non-Speech Audio in the Practice Section 41

Figure 11. Non-Speech Audio in Knowledge Check Activities 42

Figure 12. An Example of Text Equivalent 43

Figure 13. Matching Activity 44

Figure 14. Six scales of User Experience 46

Figure 15. The SDT Model of Motivation 47

Figure 16. An Example of Web Data Collected at the Backend......................................50

Figure 17. Character Door in the Practice Section...84

Figure 18. Character Child in the Practice Section 85

Figure 19. Character Woman in the Practice Section 86

Figure 20. Character Horse in the Practice Section..87

Figure 21. Character Man in the Practice Section 88

Figure 22. Character Water in the Practice Section 89

Figure 23. Character Sun in the Practice Section 90

Figure 24. SPSS Output of Box's M 96

Figure 25. SPSS Output of Bartlett's Test of Sphericity3 96

Figure 26. SPSS Output of Multivariate Test 97

Figure 27. SPSS Output of Levene's Test 98

Figure 28. SPSS Output of Univariate ANOVA Tests 99

LIST OF TABLES

Page

Table 1. Distributions of Participants 33

Table 2. An Example of the Items 49

Table 3. Descriptive Statistics of Learner Motivation 55

Table 4. Descriptive Statistics of Online Learning Experience 56

Table 5. Descriptive Statistics of Time 56

Table 6. Multivariate Test 58

Table 7. Tests of Between-Subject Effects 59

Table 8. Interaction Effect 59

CHAPTER I

INTRODUCTION

In the past several decades, researchers and practitioners sought to incorporate auditory information in learning materials (Domagk et al., 2010). They believe that effective learning occurs when people build mental representations by combing visual and auditory information (Mayer, 2005; Mayer & Betrancourt, 2005; Mayer et al., 2001; Mayer & Moreno, 2003). Several design principles have been proposed to guide the design of multimodal instruction. Examples of such principles are the redundancy principle, the modality principle, and the multimedia principle. These design principles address the importance of auditory representations that allow information to be processed in a separate auditory channel to avoid mental overload (Mayer et al., 2001).

The auditory representations discussed in these principles refer primarily to speech audio (i.e. the spoken words). However, auditory representations are not limited to speech audio, but include unaltered or slightly altered sounds such as human speech, music, environmental sounds, and electronically created musical notes (Csapó & Wersényi, 2013; Wersényi, 2010). Despite the fact that non-speech audio is an important type of auditory representation, little has been mentioned in relation to multimedia learning (Bajaj et al., 2015; Bishop, Amankwatia & Cates, 2008; Bishop & Sonnenschein, 2012).

As one neglected aspect of auditory information in current learning materials, non-speech audio plays a fundamental role in everyday life. We interpret the world around us and determine our actions according to the non-speech sounds we receive. For example, we recognize an emergency when there is a fire alarm or when we hear the sirens of a police car, even if it is out of sight. In addition, we could detect the position of the police car and decide whether to yield to the police car or not. Non-speech sound enables us to function effectively in daily life as a basic source of information. Therefore, if applying non-speech audio in the virtual learning environments, the skillset we have built up from everyday life could be applied to help better interact with learning environment, thereby leads to a better learning performance.

1.1 Statement of the Problem

Although being largely ignored in the field of education, the use of non-speech audio is valued and widely used to aid interactions in other fields. A well-known example is the use of non-speech audio in video games. To a large extent, the game players rely on the non-speech audio to navigate fast interactions with the game system. Players' performance is believed to drop when the non-speech audio is removed from the game (Jørgensen, 2009). In many computer systems, non-speech audio is used to validate and support user action. For example, users will hear a "crash" sound when a file is placed into the "trash can", a "glitch" sound is heard when an error occurs, and a "ding" sound is used to confirm the completion of a file transfer. The relationship between a non-speech sound and an action simulates the user's interactive experience with real-life objects and associated sounds (Brewster, 2002; Rocchesso et al., 2013). In addition to improving interactive user experience, the use of functional non-speech audio has the potential to influence performance (Buxton et al., 1994). As mentioned earlier, game players' performance has been observed to drop when playing without non-speech audio. The information conveyed through the use of non-speech audio in the game environments enables the game players to perform effectively. For example, non-speech audio reinforces visual represented information and enable players to monitor background processes, locate critical information, and collect environmental information (Jørgensen, 2009). All the information can be processed by the players without deviating from their primary visual tasks.

While designs in other fields often emphasize sound properties such as timbre and pitch, researchers have pointed out several reasons why the use of non-speech audio should be considered in user experience design (Brewster, 2002, Buxton, 2010; Buxton et al., 1994). These reasons align well with the design principles applied in the development of multimedia learning environments as recommended by cognitive researchers in the field of education. First, the use of non-speech audio help to reduce the amount of information needed on the screen to effectively convey information. This is particularly helpful in the use of devices with small visual displays. Moreover, with less visual information being displayed, non-speech audio help to reduce the users' visual load. In addition, when the information is presented on large screen graphical interfaces, intensive use of the user's visual system would lead to critical information being overlooked. The user's visual load could be reduced when some information is presented as non-speech audio. For example, users could benefit from auditory feedback instead of visual feedback. Therefore, the investigation is needed on how to design and apply non-speech audio for learning purposes in the field of education.

1.2 Theoretical Framework

The application of non-speech sound in learning environments is a challenging area of study. In our daily life, non-speech sound is a critical information source to guide our actions when we interact with the environment and build our understanding of the surroundings. In other words, non-speech sound is an important element generated from the interaction, meanwhile, it impacts on the interaction process. Learning through the interactions with the environments are supported by experiential learning theory and interactivity principle. In addition, the roles of non-speech audio in the interaction process resemble the Stimulus and Response theory. Therefore, guided by experiential learning theory, interactivity principle, and the Stimulus and Response theory, interaction with the surrounding environment is the basis of the application and design of non-speech audio in the online learning environments in this study. These three supporting theories are described in more details in Chapter II.

1.3 Purpose of the Study

In light of the aforementioned potential benefits associated with the use of non-speech audio in learning environments, the present study seeks to address an important gap in the literature and to provide an understanding of the effects of the use of non-speech audio in education. Given that new technologies are increasingly interactive, it is important to understand how the use of non-speech audio could benefit learning. However, without an in-depth and well-grounded understanding of the effective use of non-speech audio for learning purposes, researchers and practitioners in education tend to rely on their own constructed approach when attempting to integrate non-speech audio in learning environments, also without the support of research or empirical evidence.

This dissertation applies experiences and practices from the field of human-computer interaction to educational settings and explores the process by which educators design, develop, and implement non-speech audio in interactive learning environments. The following questions guide the investigation of this dissertation:

1. What have research studies been done on the use of the non-speech audio in the field of education? Based on these studies, what do we know about the ways in which non-speech audio have been designed and implemented to date? What needs to be done to advance scientific research in the field?

2. What are the main effects of the non-speech audio when embedded in the interaction process in the virtual learning environment?
3. Does the effect of non-speech audio change depending on the learners' characteristics?

1.4 Significance of the Study

The significance of this study lies in its attempt at investigating the effects of non-speech audio in interactive learning environments. Increasingly today, emerging technologies tend to focus on learning environments where learners have unprecedented freedom to navigate through the content (Mihalca & Miclea, 2007). Learning processes in such interactive environments simulate ways in which learners perceive and experience their surroundings in the real world. As a basic source of information, non-speech audio is essential to guide our actions in daily life, yet it is underrated as a format of information representation in the field of education.

This dissertation aims to make contributions to the educational field in several ways: 1) offer considerations for the integration of non-speech audio in the design of interactive learning environments; 2) provide design guidelines for the practical and effective use of non-speech audio; 3) improve the understanding of the effective use of non-speech audio for educational purposes with empirical evidence; 4) The findings of this study may apply to the future design and development of non-speech audio in interactive learning environments; 5) The study results might bridge conceptual gaps associated with the use of non-speech audio in the field of education.

1.5 Organization of the Study

This dissertation includes five chapters. Chapter I briefly reviews the theoretical foundations that drive research done on the use of non-speech audio used in learning. Related terminology and corresponding definitions are also described. Chapter II is a systematic literature review examining existing studies to understand the circumstances in which non-speech audio is being used in learning. Chapter III provides the research design, methodology, and procedures used to conduct this study. It also presents the population and study participants and instrumentation. In addition, it provides the details of data collection as well as the data analysis strategy. Chapter IV reports the findings of a 2 x 3 MANOVA that presents the results of self-reported questionnaires and web data collected in the experiment. Chapter V answers the research questions and the results of the investigation on how the proposed design of non-speech audio affects learner experience and performance. In addition, it also includes the discussion of the findings, the limitations of the current study, implications for the educational practices, and the recommendations for future research.

1.6 Definition of Terms

The terms used in this study are defined as following:

- Non-Speech Audio: refers to the programmed non-verbal sounds that are used in response to learners' interactions with online learning environment.
- Interaction: refers to the physical interaction between learners and the online learning environment. The online learning environment responds to learners' actions by sending the signals (Domagk et al., 2010).
- Linear Online Learning Environment: An online learning environment where limited or none interactions involved between the learners and the online learning environment.
- Interactive Online Learning Environment: An online learning environment where learners construct their understanding and knowledge through intensive interactions with the learning environment.

CHAPTER II

LITERATURE REVIEW

2.1 Theoretical Framework

Experientialists believe that effective learning is seen when placing experience at the center of the learning process (Kolb, 1984; Kolb, 2014; Mainemelis, Kolb & Boyatzis, 2002). One of the most influential experientialists, David Kolb (1984), defines experiential learning as a process “whereby knowledge is created through the transformation of experience” (p.41). Based on Experiential Learning Theory (ELT), fully engaged learners go through four main stages in the learning process. In the first stage, learners acquire knowledge from their concrete experience by interacting with the world through their senses, feelings, and emotional impulses. In the second stage, learners reflect on their experience from different perspectives in order to initiate the third stage of abstract conceptualization. In the third stage, learners generate theories and conclusions through observation and reflection. In the last stage, learners test their theories and use these theories to solve problems (Kolb, 1984). These four-stages form a continuous cycle where concrete experience functions as driving forces in the learning process.

According to experiential learning theory, concrete experiences are derived from internal and/or external stimuli. External stimuli are perceived through the learners’ senses such as touch, smell, hearing, sight, and taste, while internal stimuli are feelings and emotional impulses such as excitement, anxiety, fear, hurt, and hope (Carver, 1996). Learners gain knowledge from concrete experiences by receiving, interpreting, and responding to these stimuli during the learning process (Baasanjav, 2013; Beard et al., 2007; Carver et al., 2007). When moving to an interactive virtual environment such as digital games and simulations, digital media simulate a ‘real experience’ in which learners build their own concrete experience by interacting with the virtual world. These concrete experiences are the integration of everyday experiences and communication technology (Baasanjav, 2013; Murphrey, 2010; Riedel et al., 2007; Trevitte & Eskow, 2007). In other words, in an interactive virtual environment, learners receive, interpret, and respond to stimuli from the virtual world during the learning process.

Similar to the experientialists, cognitive multimedia learning theorists also consider the interaction between learners and virtual learning environments to be beneficial for learning. Such interaction is defined as learner-environment interactivity that refers to a “reciprocal activity between a learner and a multimedia learning system, in which the [re]action of the learner is dependent up the [re]action of the system and vice versa” (Domagk et al., 2010, p1025). Once registered by sensory signals, the learners are expected to initiate an action in respond to the signals. The virtual learning environment immediately provides feedback to the learners according to which the learners organize and integrate information and then store knowledge for future use (Aleksy et al., 2016; Borchers, 2000). When learners interact actively with a virtual learning environment, that active engagement allows deep learning to occur. This view of learner-environment interactivity stems from a human-computer interaction perspective (Markus, 1987; McMillan, 2002; Moreno & Mayer, 2007).

A major point of agreement between experientialists and cognitive multimedia learning theorists is that interaction between learners and virtual learning environments is essential for effective learning (Domagk et al, 2010; Markus, 1987; McMillan, 2002; Kolb, 1984). These theorists believe, in order for learning to take place, learners need to be first attracted by some signals in the learning environment, and then actions occur as a result of the learner’s respond to the signals. This interaction pattern aligns with the Stimulus-Response (S-R) theory recognized in behaviorism (Toates, 1997). A stimulus is assumed to have a certain strength of tendency to produce a response. Given immediate feedback, learner behavior could be modified and the desired response could occur. As argued by Toates (1997), S-R theory is valid, particularly in shaping learning behaviors. Accordingly, when S-R theory is integrated within a cognitive framework, it changes learner behavior as well as their learning experiences (Bindra, 1978; Epstein, 1982; Fentress, 1976; Toates, 1997).

As an essential element in a learning process, interactions between learners and virtual learning environments have made it possible for the use of non-speech audio to play an important role in the learning process. In the process of each interaction, non-speech audio can be implemented by itself or together with other input formats to function as a stimulus to trigger an action or as a feedback to reinforce an action (Saffer, 2013). When we take a granular view (see figure 1), we see the entire interaction process is first triggered by a stimulus (visually or aurally, or both), depending on the complexity of the interactive learning environment, non-speech audio can be used alone or co-present with other input formats to trigger learner action. Next, the learner will respond to the stimulus by acting on the interactive learning environment (Learner Action). And then, when the learner's action is considered to be a correct response, non-speech audio can be used alone or co-present with other input formats to confirm the learner's correct action (Feedback). When the learner's action is considered to be an incorrect response, non-speech audio can be used alone or co-present with other input formats to modify and guide the learner's next action until the correct response is received by the learning environment (Aleksy et al., 2016; Franinović & Salter, 2013).

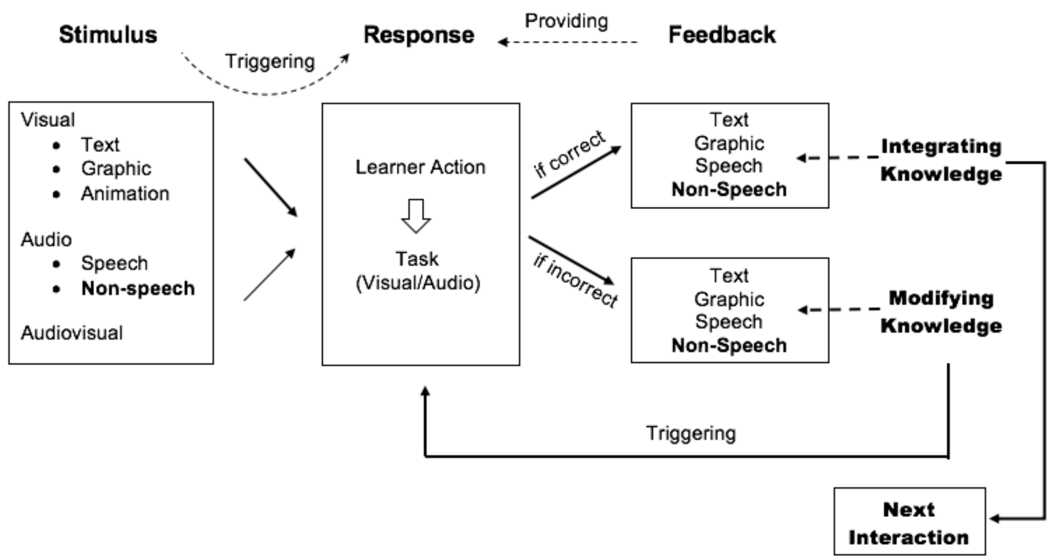


Figure 1. Conceptual Model of Non-Speech Audio Design in the Interaction

Based on the above conceptual model, different types of information can be carried by non-speech audio, including the following four broad categories: 1) Alert: to alert the learner to pay an immediate attention to something that occurred in the environment. In general, little information is carried in alert sound in terms of the details of the event or task (Buxton, 1989; Walker, Nance & Lindsay, 2006). For example, a beep simply indicates that a problem is detected and requires immediate attention. The basic beep sound does not indicate the nature or the cause of the problem. 2) Confirmation: to confirm an action. In this case, an expectation comes naturally from the learner's real-life experience and the learner expects the system to confirm each action taken by the learner as the learner interacts with objects within the learning environment (Brown, Newsome, & Glinert, 1989). For example, in a navigational interaction, when clicking on a button or on a hyperlink, the learner will hear a click. 3) Understanding: to understand the result of the action. In addition to confirming that their action has been executed, learners expect to know the result of their actions, whether they respond to the input in a correct way or not. Non-speech audio sends positive or negative feedback to the learners and enable them to understand the result of their actions (Kramer et al., 1999). For example, if you want to download an app in the Apple store, a “ding” sound combined with a visual check mark indicates the successful installation of the app after hitting the “Get” button. 4) Modification: to modify and guide an action. When non-speech audio is designed to permit data exploration, it will serve as an indicator to modify and guide learner action. For example, in a target tracking task, non-speech audio could be a helpful constraint on actions or a directive to guide learners’ movement through dynamic interactions (Dyer, Stapleton, & Rodger, 2017).

The integration of non-speech audio in virtual learning environments can lead to enhanced learning. Non-speech audio has great potential for augmenting the concrete experience in virtual learning environments. First, it provides an opportunity to combine the advantage of external stimuli including visual, auditory, and tactile senses in the interaction process. Such multisensory experience simulates how a person interacts with everyday tasks in real-life, and consequently immerses the learners in virtual learning environments during the interaction process (Buxton et al., 1994; Buxton, 2010; Königschulte, 2015; Rocchesso, Serafin, & Rinott, 2013). Second, it brings the possibilities to stimulate motivation in learning by mediating the internal stimuli such as excitement, anxiety, and fear (Danna et al., 2015; Königschulte, 2015; Schaffert et al., 2009). A widely known example of using non-speech sound to arouse feelings and emotions is the film industry. Non-speech sound plays a major role to highlight critical moments and enhance realism in movies. A good sound design is crucial for attracting audiences into watching a movie (Kerner, 1989).

In addition to the enhancement of concrete experience, non-speech audio has significant potential for improving learning performance (Danna et al., 2015; Jørgensen, 2009; Molholm et al., 2002). When co-presenting with visual input as a stimulus in a complex interactive environment, non-speech audio could help learners to enhance stimulus detection by reducing perceptual ambiguity (Molholm et al., 2002). When functioning as feedback, non-speech audio enables learners to stay focused and prevents them from deviating from their main visual tasks (Berman & Colby, 2002; Brown et al., 1989; Calandra et al., 2008; Daniels, 1995; Daniel & Moore, 2000; Hativa et al., 1987). While performing a main visual task, any additional feedback presented to the learners may be processed more efficiently if it requires a different type of processing (Spence, Ranson, & Driver, 2000). Simply stated, it is more efficient to provide feedback through an auditory channel when learners are required to perform a main visual task. Furthermore, it is efficient to use non-speech audio to provide feedback to modify and guide learner's movement, particularly in a dynamic interaction process (Danna et al., 2015; Dyer, Stapleton, & Rodger, 2017; Molholm et al., 2002).

In summary, the addition of non-speech audio in the interaction process is believed, according to the aforementioned conceptual model, to enhance the learner's online experience, stimulate learner motivation, and encourage learners to engage in learning, thereby leads to better learning performance. To effectively apply non-speech audio in virtual learning environments, the first step will be to develop an understanding of the circumstances in which non-speech audio is being used in learning. Therefore, this chapter provides a systematic overview of previous research on the use of non-speech audio in education. The theoretical framework of this study is grounded in experiential learning theory, learner-environment interactivity, and stimulus-response theory, and hence this framework serves as a lens through which to review the use of non-speech audio in existent literature.

2.2 Systematic Review Procedure

In order to summarize available research on the impact of non-speech audio on learning, this study examines the literature in a systematic way. Given the fact that this review attempts to explore the use of non-speech audio in education, the chosen articles were taken from major educational databases such as ERIC, PsycINFO, and Academic Search Complete (ASC).

2.2.1 Search Criteria

Initially, broad search terms such as sound, audio*, music, learning, performance, education, and training were applied to locate relevant articles. All possible combinations of these search terms were applied to these databases to identify the most relevant studies. After a couple of rough screenings, 10 research studies that examined the use of non-speech audio in education were identified. During an analysis of these 10 research articles, I extracted several keywords from the articles as well as from the database descriptors. I then applied different combinations of these keywords and descriptors to all of the educational databases and also created custom search terms for different databases to locate the most relevant studies. For example, in Eric, a combination of search terms related to audio (i.e., sound effects, audio*, audio stimuli*, audio cue*, and noise) and learning outcomes (i.e., attention, retention, comprehension, and attainment) were used. (See Appendix A for detailed information on search terms used in all databases.)

2.2.2 Inclusion and Exclusion Criteria

Once the articles were pulled from the databases, these documents were screened in two phases: abstract screening and full text screening. During screening, articles that focused on the use of non-speech audio and described how non-speech audio could be used to affect learning were selected. All selected articles have been published in peer-reviewed English language journals.

The use of the identified search terms yielded a total of 849 articles from the databases. There were 489 articles from Eric, 127 from PsycINFO, and 233 from ASC. From these 849 articles, 49 duplicates were removed. Next, during the abstract screening phase, the abstracts of the 800 remaining articles were examined to determine relevancy to the issues on which this review focused. First, some articles had the word “sound” in their titles, but it was used as an adjective. A large number of articles were eliminated in this phase because they did not meet the criteria for this literature review and did not focus on the use of non-speech audio for learning. For example, some articles explored the effects of speech audio on learning by comparing written instruction to auditory instruction. Other articles investigated the use of written feedback versus auditory feedback. The use of speech audio as cues were also covered by some researchers, while others examined the use of stand-alone graphics in contrast to graphics used together with narrative explanation. Some articles described studies that used speech audio to teach music education or foreign languages. Still more articles were eliminated because they assessed the effects of multimedia on learning. Other articles investigated audio-based computer learning courses, but the focus was on the effectiveness of these programs. Upon the completion of the abstract screening phase, a total of 42 articles proved to be relevant to the focus of this review: 21 from Eric, 9 from PsyINFO, and 12 from ASC. The references within each of these articles were then examined. Based on this information, 3 additional articles were identified, bringing the number of articles available for review in the second screening phase to a total of 45.

In the second phase of the review, the full text of these 45 articles were examined. The same criteria as used in the abstract screening phase, were also applied in this full text screening phase. A total of 27 articles were excluded during this second screening phase because they did not meet these criteria. For example, 4 articles were removed because they focused on environmental noise. One study published by Calvert (1981) focused on the use of sound to measure students' selective attention. Another study published by Mayer (2005), addressed multimodal learning through the use of multimedia. The multimedia used in this experiment included text, graphics, and audio, but the focus of Mayer's study was on the difference between the use of text with audio and graphics with audio. Five more studies focused on responses to visual and auditory signals rather than on the use of non-speech audio to influence student learning (Guttentag, 1985; Ho, 2005). Other articles were removed because they focused on techniques used for the design of non-speech audio. As a result of these deductions, a total of 18 articles were left to be evaluated in this literature review.

2.3 Results

The 18 identified articles which focused on the use of different types of non-speech audio in education were published between 1979 and 2016. The sample size of these studies ranges from 20 to 120 participants. All 18 studies cover grade levels spanning K-12, special education, and higher education, as well as corporate and professional training. Moreover, these studies investigated the use of non-speech audio across diverse instructional programs and materials including in-class presentations, instructional software, computer simulations, auditory stories, television stories, and computer-based programs. While all studies examined the use of non-speech audio in learning, they use different terms according to the function of non-speech audio used in the studies. For example, background music (n=6), auditory cues (n=3), and sonification (n=1) (see figure 2 for more details).

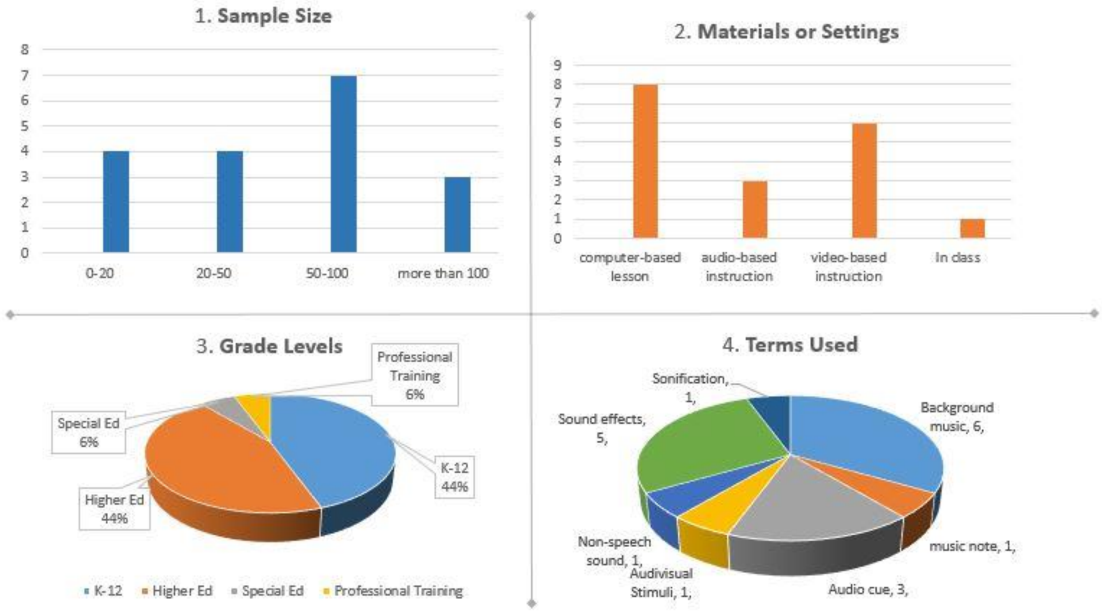


Figure 2. Result of Systematic Review

In the 18 reviewed studies, non-speech audio was primarily identified for three primary uses. Non-speech audio was used to: (1) serve as a stimulus to direct learners' attention to key concepts or critical scenes (Bajaj, 2015; Bishop, 2000; Boltz, 1991; Calvert, 1985&85&87&89; Hativa, 1987; Hupp & Allen, 2005; Mann, 1979), (2) represent information (Ferati, 2012; Pfeiffer, 2008; Vickers, 2002), (3) create atmosphere and moods to engage learners and enhance the quality of the learning experience (Chou, 2010; Fumham, 1997; Grice, 2009; Perham, 2011&2014;).

2.3.1 Primary Use One - Serves as Auditory Stimuli to Trigger Attention

The first and foremost of the 3 above-listed primary uses of non-speech audio in learning environments is to function as auditory stimuli to direct learners' attention in the learning process. Different types of non-speech sound, including environmental sound, musical notes, and instrumental music, were used in 10 studies as auditory stimuli that were closely associated with key concepts or critical scenes in videos, auditory story clips, or computer-based lessons (Bajaj, 2015; Bishop, 2000; Boltz, 1991; Calvert, 1985&85&87&89; Hativa, 1987; Hupp & Allen, 2005; Mann, 1979).

Three studies used environmental sounds as auditory stimuli (Bajaj, 2015; Mann, 1979; Bishop, 2000). For example, the sound of water flowing through a tap was presented together with a narration that described how to add water to a chemical compound (Bajaj, 2015); clinking dishes and muffled conversations along with visual images to cue student attention to certain areas in presented images of a restaurant (Bishop, 2000). Seven studies used musical notes to mark important events (Bajaj, 2015; Calvert, 1985&85&87&89; Hativa, 1987; Hupp & Allen, 2005). These musical notes were random brief notes that were not directly related to the event with which it was associated. For example, the sound of a slide whistle was used in videos to direct student’s attention to critical scenes (Calvert, 1985&85&87&89). A burp sound was associated with the key concept of acidity in a video instruction (Bajaj, 2015). One study (Boltz, 1991) associated instrumental music with critical scenes to trigger students’ attention.

Nine studies used either environmental sounds, musical notes, or instrumental music to cue the student to critical events. Only Bajaj (2015) used a combination of environmental sounds and musical notes as auditory stimuli in his study. In addition, eight studies tested the effects of non-speech audio as auditory stimuli in instructional videos and auditory productions, and two studies examined non-speech audio used in computer-based lessons (Bishop, 2000; Hativa, 1987).

2.3.2 The Effects of Non-Speech Audio as Auditory Stimuli

Ten studies used environmental sound and musical notes as auditory stimuli to direct attention to important information presented in videos, auditory clips, and computer-based lessons. Of these ten studies, nine studies examined the effects of non-speech audio on retention (Bajaj, 2015; Boltz, 1991; Calvert, 1985&85&87&89; Hativa, 1987; Mann, 1979), comprehension (Calvert, 1985&85&87&89; Hativa, 1987; Mann, 1979), and completion rate of training programs (Hupp & Allen, 2005). One study explored the effects of non-speech audio on motivating learners to engage in computer-based lessons (Bishop, 2000).

2.3.2.1 Retention

Seven studies examined the impact of non-speech auditory stimuli on knowledge retention. All seven studies suggested that non-speech auditory stimuli helped learners improve their scores on short-term knowledge retention tests (Bajaj, 2015; Calvert, 1985&85&87&89; Hativa, 1987; Mann, 1979). Additionally, Bajaj (2015) and Mann (1979) also observed positive effects associated with non-speech auditory stimuli on long-term knowledge retention in their studies in which the results showed that students in the experimental group that received auditory stimuli outperformed the students in the control group that received no auditory stimuli. According to the findings of these seven studies, auditory stimuli reinforced the key information in the instruction, making it easier for learners to remember the information.

Calvert (1985&85&87&89) conducted a series of experiments to assess the effectiveness of non-speech auditory stimuli on retention. In these studies, auditory stimuli were embedded in instructional videos that were slow-paced, regularly-paced, fast-paced, and variably-paced. Findings indicated that non-speech auditory stimuli were particularly effective in helping students recognize more items in fast-paced videos. The authors suggested that this difference could be because the events in the high-paced videos occurred so fast that the students needed guidance to direct their attention to important events (Calvert, 1985&85&87&89).

In Boltz's study, instrumental music was used as auditory stimuli to highlight important events in a video recording (Boltz, 1991). He discovered, when the mood of the music was congruent with the content of the video, auditory stimuli were effective to help learners retain highlighted concepts. Conversely, when the mood of the music was incongruent with the video scenes, learners performed even worse than when no music was used at all.

2.3.2.2 Comprehension

Five studies assessed the effects of non-speech auditory stimuli on learners' performance on comprehension tests. According to the results presented in these studies, non-speech auditory stimuli helped students improve their scores on comprehension tests (Calvert, 1985&85&87; Hativa, 1987; Mann, 1979). In Hativa's study, comprehension tests were conducted at two distinct time points: one was run immediately after learning, and the other was administered one month after the experimental treatment. The results indicated that the positive effect of non-speech auditory stimuli on students' comprehension performance remained after one month. Moreover, Hativa discovered that the use of non-speech auditory stimuli was particularly beneficial for the learning comprehension of low aptitude students.

2.3.2.3 Completion Rate

Hupp and Allen (2005) conducted a case study to observe how the use of non-speech auditory stimuli influenced on task completion rate. His study focused on a learning experience designed to improve child-parent interactions. Participants were required to make critical statements at different times throughout the training. The dropout rate was high in previous training because the participants were unable to make the required number of statements. In

Hupp's case study, a beep tone was played to remind the participants to make statements at particular moments in the first half of the training. The auditory cues were gradually removed at the second half of the training, but the participants were able to complete the training and even made more statements than it was required.

2.3.2.4 Engagement

Bishop (2000) performed a case study to observe how effective non-speech auditory stimuli were on engaging 4 learners in a computer-based lesson in a school lab. When taking the lesson without any non-speech audio, unmotivated student F worked through the lesson within seven minutes, poorly focused student D stayed for seven minutes on the introduction, timid student C managed to complete the lesson but he felt it was meaningless, and eager student A also completed the lesson but was frustrated because of some mistakes she made. A week later, the same 4 participants took the same lesson again, but this time with a sound enhanced environment where environmental sounds were co-presented along with visual images. For example, the sounds of crickets chirping, a constant dry wind blowing, and a wolf howling in the distance were heard while the visual image of a desert was presented on the screen. As background music, an ominous orchestral theme played while the lesson's logo "Thinking Zone" scrolled across the screen. According to Bishop, even though they were presented with the same lesson for a second time, the 4 students were more engaged with the content when non-speech sounds were used. In particular, unmotivated F was anxious to explore the scene and remained interested even after he completed the lesson. Poorly focused D actively explored the lessons and requested more information about the content after finishing the lesson.

2.3.3 Primary Use Two – Representing Information

The second primary use of non-speech audio in learning environments is to represent information. Three studies examined using non-speech audio to represent instructional content: 2 used non-speech audio alone in the classroom (Ferati, 2012; Vickers, 2002) and one study used non-speech audio together with graphics in computer-based lessons (Pfeiffer, 2008). Ferati (2012) focused on visually impaired learners, while Pfeiffer (2008) and Vickers (2002) investigated how regular learners learned when instruction was communicated through musical notes.

Ferati (2012) and Pfeiffer (2002) used only non-speech audio to teach different concepts. Ferati (2012) used environmental sound to describe the main concepts covered in learning. For example, the sound of a hammer hitting a nail was used to represent 3 related concepts: hammer and nail, building construction, and designing architecture; the sounds of keys jangling and a car engine starting, the related concepts were keys and car, and the associated concepts were travel and security. Pfeiffer (2002) used musical notes in combination with graphics to represent the categorical information. For example, a buzzy sound with a curved line represented sheep, a piercing sound with straight line referred to rodents, and an “ooo” sound with a dotted line pointed to coyotes. These musical notes were created through Musical Instrumental Digital Interface (MIDI) software. In Vickers’ study, musical notes were designed to represent different computer programming concepts. For example, the sound of a fog-horn represented the concept of IF and IF….ELSE while the sound of piano notes “dah dah dit” referred to the concept of FOR (Vickers, 2002). These non-speech sounds had the same concrete meaning as written text.

2.3.4 The Effects of Information Represented Through Non-Speech Audio

As previously discussed, a study conducted by Ferati (2012) focused on the use of non-speech audio to communicate critical information to visually impaired students in a special education school. The participants were expected to study 35 sets of environmental sound for two weeks. Each environmental sound represented three related concepts. And then, they were given three tests at three different times: 1 week, 4 weeks, and 7 weeks after the training. The findings indicated that the use of these non-speech sound was highly effective in assisting visually impaired students to recognize and remember different concepts. In addition, female students had a more accurate recognition of non-speech audio with associated concepts than male students.

Also mentioned earlier on, 2 studies conducted by Pfeiffer (2008) and Vickers (2002) focused on learners with normal vision. Both studies used MIDI software to create musical notes to communicate important information. Vickers discovered that all participants were able to identify 13 out of 14 constructs communicated by musical notes. Only one construct seemed to be difficult to identify, most likely because of its complexity (FOR..TO was nested within REPEAT), as explained by Vickers (2002). In Pfeiffer's study, results indicated no significant difference in test scores between the control group that studied categorical information from the regular graphics and the treatment group that studied from the sonified graphics. However, the use of musical notes did have a positive effect on the online learning experience (Pfeiffer, 2008).

2.3.5 Primary Use Three – Create Atmosphere and Moods

The third primary use of non-speech audio is as background sound to enhance the quality of the learning experience through the creation of atmosphere and moods. Five studies used music as background in computer-based lessons (Chou, 2010; Fumham, 1997; Grice, 2009;

Perham, 2011&2014). Specifically, music played in the background when students performed a cognitive task such as reading.

2.3.6 The Effects of Background Audio on Learning

Five studies examined the use of non-speech audio (music) as background to enhance learning experience (Chou, 2010; Fumham, 1997; Grice, 2009; Perham, 2011&2014). The non-speech audio used in these studies were background music playing in computer-based lessons for entertainment purposes. According to the findings from these 5 studies, the learners were negatively affected by all styles of background music including classic, hip hop music, and instrumental music. Effects were particularly negative when students were performing cognitive tasks such as reading (Chou, 2010; Fumham, 1997; Grice, 2009; Perham, 2011&2014). In addition, Perham (2014) found that both liked and disliked music were equally distracting. He also pointed out that students reading with liked and disliked music that included lyrics performed worse than those who were exposed to non-lyrical music or no background music. Fumham (1997) detected learners with introverted and extroverted personalities were equally distracted by background music while performing cognitive tasks.

2.4 Discussion

In the past decades, research seems to pay little attention to the use of non-speech audio. While a large number of articles pulled from educational databases during the search, most of them focused on the use of speech audio for learning. Unlike an extensive investigation of the effects of speech audio, only 18 articles examined the impact of the use of non-speech audio on learning in the past 30 years. Although the numbers of studies are limited, these studies still help to build an understanding of how non-speech audio is being used in education. In addition, the findings from the studies provide useful guidelines for creating and applying non-speech audio for learning purposes.

According to the existent literature, non-speech audio is most often used as auditory stimuli. When serving as auditory stimuli, non-speech audio is believed to facilitate learning. First, non-speech auditory stimuli have a positive effect on memory retrieval. Auditory stimuli are typically embedded around key concepts in instructional videos, auditor story clips, and computer-based lessons. During recall, auditory stimuli could serve as a memory anchor to retrieve content and evoke associated information (Calvert, 1985&85&87&89; Bajaj, 2015; Hativa, 1987; Mann, 1979). Second, the use of non-speech auditory stimuli affects the way learning information is processed and retained. A systematic application of auditory stimuli directs learners' attention to specific information, emphasizes key concepts, and organizes learning information (Bishop, 2000; Hativa, 1987; Calvert, 1985&85&87). If using non-speech audio is used to build information structures, learners are able to transfer knowledge to new learning situations more effectively (Bishop, 2000&2008). Third, non-speech audio, and environmental sound in particular, adds a sense of immersion to learning experiences. For narrative audio clips and instructional videos, associating environmental sound to critical scenes helps to visualize the message and keep learners motivated during the learning process (Bishop, 2000; Psotka, 1995).

However, the proper use of different types of sound as stimuli is important. Environmental sound is associated with the environment and well known. Therefore, when presented together with other information representations such as text, graphics, and narrations, environmental sound is effective in cueing the related instructional message. On the other hand, a careful design is required for musical notes to be effective to cue the critical information in the instruction. Since musical notes are not naturally relevant to the presented information, selected musical notes can be effective enforcers of critical information when regularly patterned in the instruction (Bajaj, 2015; Hativa, 1987; Hupp & Allen, 2005). If learners fail to understand the connection between the musical notes and the content, these musical notes could become irrelevant and even distracting (Boltz, 1991).

In addition to serve as auditory stimuli, non-speech audio (i.e., music) is also placed as the background in some computer-based lessons. Unlike the positive effect of auditory stimuli, the use of non-speech audio in the background such as background music is considered to overload learner's information processing capacity and have a negative impact on the performance of other cognitive tasks. In other words, all styles of background music should be used with caution when learners are required to perform cognitive-related tasks. Background music could have a distracting effect and could lead to even more negative results when content irrelevant. Still, background music is best used to create atmosphere or mood during the introduction, major content transitions, or at the end of a learning experience. In the typical computer-learning environment, these three areas do not involve significant cognitive tasks and background music could add value. As seen in Bishop's computer-based lesson, background music played at the beginning creates a "cinematic style" (Bishop, 2015) that learners are motivated to explore the lessons actively.

One of the earliest uses of non-speech audio is to represent information using aural attributes such as loudness, pitch, and frequency (Buxton, Gaver & Bly, 1994). Specifically, a variety of data information can be represented by sound because of the hierarchical attributes in auditory variables (Lunney and Morrison, 1990). The use of non-speech audio to represent information has also been discussed in the reviewed literature. Non-speech audio is found to be effective for visually impaired learners in recognizing and understanding instructional message. Compared to visually impaired students, benefits are less clear for learners with normal vision. This difference might be due to the ways in which visually impaired learners and normal vision learners perceive sound (Pfeiffer, 2008). In addition, when using non-speech sound to represent information, it certainly requires training to help learners recognize patterns and structures in the information presented by sound. Although only a few studies were identified in education, researchers from other disciplines have explored the use of non-speech audio to convey information for years (Hermann & Hutt, 2011). This subject area is called sonification that is to "translate relationships in data or information into sound(s) that exploit the auditory perceptual abilities of human beings such that the data relationships are comprehensible" (Walker and Nees, 2011, p. 9). As a core component of the auditory display, sonification attempts to investigate data exploration and pattern recognition in non-speech audio. The notion of using non-speech audio to convey information has arisen in recent years, but it will not be the focus of this dissertation study.

2.5 Reference Reviewed Studies Back to the Conceptual Model

As explained in the conceptual model of non-speech audio design earlier, interaction between learners and virtual environments is essential for effective learning. The interaction has made it possible for the use of non-speech audio to play roles as the stimuli or feedback in the learning process. Meanwhile, the addition of non-speech audio in the interaction process brings the opportunity to synchronize multiple senses that augments learner online learning experience and keeps them motivated in learning.

According to the findings of the systematic review, the predominant use of non-speech audio in learning environments is to function as auditory stimuli. Different types of non-speech sound have been investigated to understand the effect of non-speech auditory stimuli in learning settings including instructional videos, auditory clips, and computer-based lessons. These findings, to some extent, have confirmed a part of the conceptual model, i.e., non-speech audio is effective in alerting learners to pay attention to something that occurs in the learning environment. However, the effect of non-speech auditory stimuli has only been examined in linear learning environments where the interaction of the learners with the learning environment either was limited to clicking buttons such Next, Previous, and Submit (computer-based lessons), or was completely not required (instructional videos and auditory clips).

According to experientialists, cognitive multimedia learning theorists, and S-R theorists, learners are expected to respond to stimuli in learning environments in order to gain experience and knowledge. In linear learning environments, learners might respond to a designated stimulus and reflect on what he/she observes. Such response is a covert cognitive process that makes it difficult to identify whether or not the desired response has occurred. Knowledge checks or assessment are usually given at the end of each learning segment to verify learners' understanding of the instruction (Hattie and Timperley, 2007). Consequently, the role of non-speech audio is relatively limited and most often used as the stimuli to direct the attention in research as well as in practice (see figure 3).

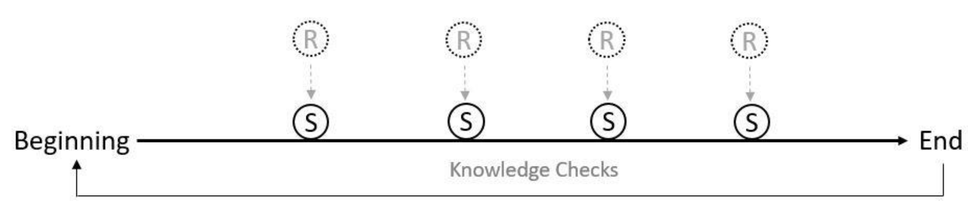


Figure 3. One Learning Segment in Linear Learning Environment

Current learning environments become increasingly interactive and complex, yet the information is still mostly presented through text, graphics, and animations that rely intensively on the visual system (Brewster, 1997; Bajaj et.al, 2015). In this case, visual stimuli could be missed due to the overloaded visual system (Brewster, 1997). Non-speech auditory stimuli can direct learner attention and reduce demands on the visual system. Moreover, non-speech audio could be more effective as a stimulus in the interactive learning environment than it is in the linear learning environment. Calvert (1989) discovered that non-speech auditory stimuli are particularly effective in helping learners locate the key information in fast-paced instructional videos. In a high-paced learning environment, the events occur so fast that the learners need guidance to direct their attention to the key information.

Based on the conceptual model, non-speech audio can be used to provide feedback on learners' action during the interaction process. While performing a main visual task, any additional information presented to the learners may be processed more efficiently if it requires a different type of processing (Spence, Ranson, & Driver, 2000). Simply stated, it is more efficient to provide additional information through an auditory channel when learners are required to perform a main visual task. Additionally, the use of non-speech audio to deliver feedback messages does not add to the learners' cognitive load (Brewster, 1997).

Moreover, as a basic source of feedback information, we constantly depend on non-speech auditory feedback to guide our actions when performing everyday tasks. For example, when driving a car, we constantly synthesize information transmitted through visual and auditory feedback to decide when to accelerate, stop, or yield. By adding non-speech auditory feedback, the synchronization of visual, auditory, and tactile senses simulates how we interact with everyday tasks in real-life. As a result, the integration of non-speech audio helps to augment learner online learning experience in an interactive learning environment.

2.6 Conclusion

The purpose of this review is to understand the use of non-speech audio for educational purposes. This is significant because non-speech sound is an important information provider that gives complementary information about learning environments. The combination of visual and auditory information in an interactive virtual environment is a powerful tool for interaction (Brewster, 2002). However, non-speech sounds, as an important part of auditory information, have been rarely implemented in learning materials (Bishop, 2008, Calandra, 2008). Moreover, the numbers of studies that do focus on the use of non-speech audio for learning are also limited. This can be seen from this systematic review for which only 18 studies could be examined to explore the effects of non-speech audio on learning. Despite the general immaturity of non-speech audio as a field of study in education, conclusions drawn from the reviewed studies provide guidelines and recommendations to designers when consider incorporating non-speech sound in learning materials. More research and testing are needed to gain a deeper understanding of non-speech audio playing roles other than auditory stimuli. Furthermore, it is important to conduct more studies on using non-speech audio in interactive virtual learning environments.

CHAPTER III

METHODOLOGY

This research study was performed in two phases. Phase one was the research planning stage during which a blueprint of this study was drafted to guide the data collection method, instrumentation, and data analysis. The instructional program used in this research experiment was also designed, developed, and tested in this phase. Phase two was the research implementation and evaluation. This research was implemented in the spring and summer semesters in 2018. The data were collected from undergraduate students enrolled in the EPSY 433, INST 301, and EPSY 435 courses offered by the Department of Educational Psychology at Texas A&M University. The gathered data were reviewed and analyzed to understand the connections between variables, to formulate facts, and to discover recurrent patterns.

This chapter provides detailed descriptions of research activities conducted in phase one. It begins with research design, followed by the identification of the target population and participants. It also covers the development of the instructional program and data collection procedure.

3.1 Experimental Research Design

The purpose of this research study is to examine the effects of non-speech audio in interactive learning environments. A quantitative research method was selected as the most suitable method to address the research questions. By examining broad trends and the shape of the distribution in the numerical data, quantitative research methods aim at exploring the patterns, correlations, or causal relationships between variables (Creswell & Clark, 2017; Leavy, 2017).

Guided by the quantitative research paradigm, this study used an experimental design to investigate the effects of non-speech audio on learning. The purpose of using an experimental design is to control all possible independent variables except the one being manipulated (Kerlinger, 1970). The comparison of the control group to the experimental group is designed to make a change in the value of the manipulated variable and to observe the effects of that change on the dependent variables.

As depicted in figure 4, participants were randomly assigned to the experimental and control groups using a random number generator in Excel. The manipulated independent variable was the version of the online instructional program. Specifically, the experimental group received a 30-minute online instructional program with non-speech audio embedded in the interactive learning environment, whereas the control group received the same instructional program without non-speech audio. The effect of experimental manipulation was observed through the dependent variables of online learning experience and motivation of online learning. The dependent variables were measured through self-reported surveys and the data recorded at the back end of the online instructional program. The surveys were embedded at the end of the online instructional program and participants were asked to take the surveys immediately after the completion of online learning.

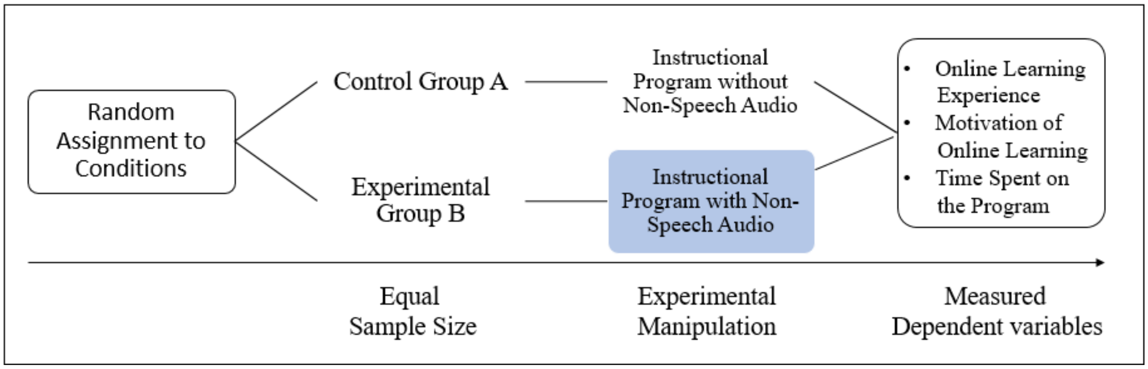


Figure 4. Experimental Research Design Process

According to the results of informal focus group interviews conducted by the researcher, learner styles and attitudes toward online learning have an impact on their online learning experience. Therefore, survey items on self-perceived learner styles, previous online learning experience, and attitude toward online learning were distributed along with other surveys.

3.2 Population and Participants

The target population for the current study consisted of Non-Chinese speaking adult learners at entry level (i.e., no prior knowledge of Chinese language). Due to the convenient accessibility, all undergraduate students (ranging from sophomore to senior students) enrolled in the Educational Psychology (INST 301), Lateral Thinking (EPSY 433), and Educational Statistics (EPSY 435) during the spring and summer semesters of 2018, received an invitation to participate in this study. In the end, 140 undergraduate students participated voluntarily in this research study.

These participants represented various visual, auditory, kinesthetic learning styles, and they had different attitudes towards online learning, ranging from neutral to positive. None of the participants were found to have a negative attitude towards online learning. The following table provides more detailed information about the distributions of the participants on perceived learning styles and attitudes towards online learning (see Appendix C for the learner characteristics questions).

Table 1. Distributions of Participants

Group	Learning Styles	Attitudes
Control Group (n = 70)	Visual = 40% Auditory = 10% Kinesthetic = 23% Combined = 27%	Negative = 0% Neutral = 26% Positive = 73%
Experimental Group (n = 70)	Visual = 44% Auditory = 17% Kinesthetic = 20% Combined = 19%	Negative = 0% Neutral = 26% Positive = 73%

As the participants were to learn and practice the fundamental strokes and the correct order of eight entry-level Chinese characters, they were required to have no prior knowledge of written Chinese language. Therefore, anyone who had any knowledge of, or were familiar with, the Chinese language were not recruited to participate in this study. Participants were randomly assigned to control and treatment groups according to the experimental design method. These participants worked through the online instruction at their preferred location, at times that were convenient to them, and using their own electronic devices.

3.3 Instructional Program – Learning Chinese

The online instructional program was a custom-designed website titled *Learning Chinese*. This instructional program allowed learners to gain knowledge of the fundamental strokes and the correct order of eight entry-level Chinese characters by interacting with the virtual learning environment. One of the main reasons for using a custom-designed website was the ability to plan and customize learner-environment interactions as needed. Moreover, the delivery of instruction through a custom-designed website eliminated the need for a Learning Management System (LMS) to host the instructional content and to track learner progress. Instead, the website was designed with a built-in flexibility at the backend that incorporated tracking codes into every page as needed to collect data on learner interactions with the learning environment.

3.3.1 Content Analysis

The subject area of this instructional program was basic Chinese characters. The learning objectives for this 30-minute Chinese character instruction were: 1. the learners will be able to identify the meaning of all eight Chinese characters; 2. the learners will be able to identify the correct stroke orders of all eight Chinese characters. The instructional content and knowledge check activities were designed in alignment with above learning objectives.

The instructional content covered a brief historical overview and practice of eight entry-level Chinese characters. The historical overview outlined the evolution of the characters. After they worked through the historical overview section, the learners understood that some Chinese characters are pictograms (see figure 5). Familiarity with character evolution could help the learners recognize some Chinese characters.

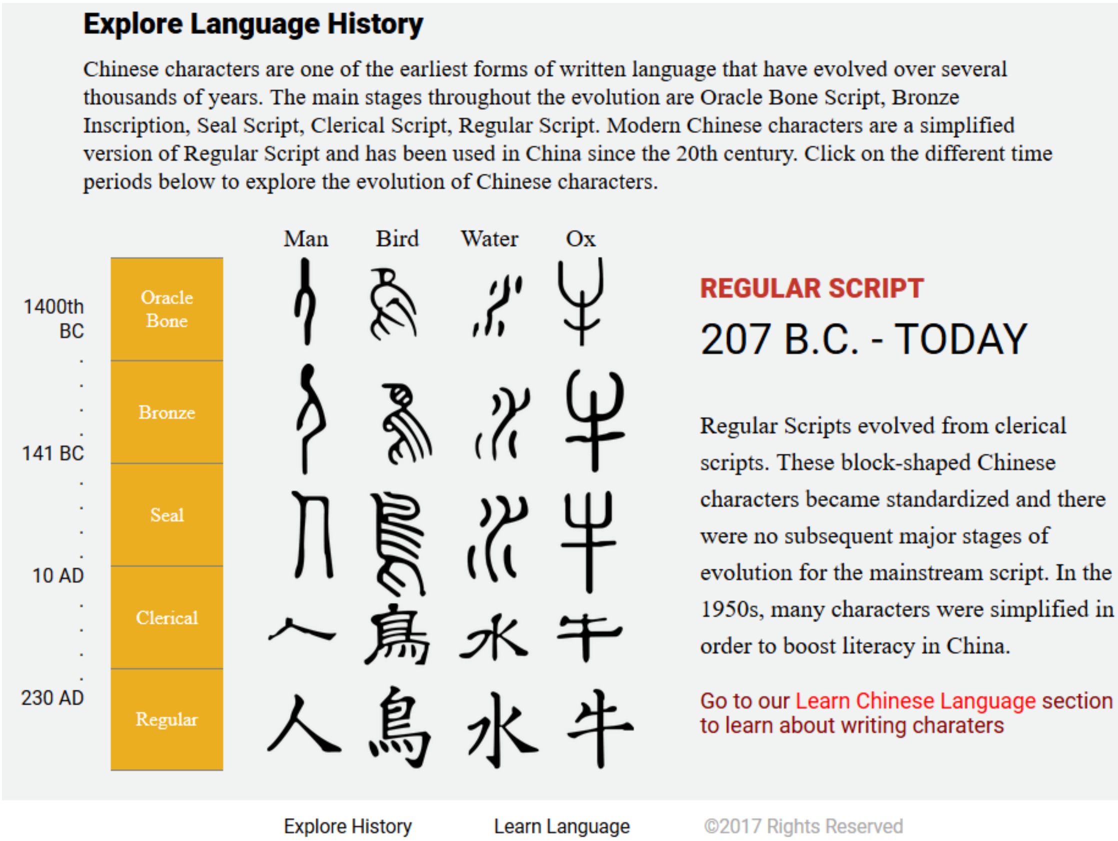


Figure 5. Instructional Content of Historical Overview

The main instructional content was the practice of entry-level Chinese characters. Pictograms were used to help the learners understand the characters they learned. Since not all Chinese characters have corresponding pictograms, eight Chinese characters which have pictograms were identified and used in this study. Based on the level of writing difficulty, these eight characters were presented in the following order in the *Learning Chinese* program: 门 (Door), 山 (Mountain), 子 (Child), 女 (Woman), 马 (Horse), 夫 (Man), 水 (Water), 日 (Sun).

The instructional strategy used to teach Chinese characters stroke orders was designed based on the conceptual model described earlier in Chapter 2. The participants constructed their own understanding and knowledge of character stroke orders through hand-on learning experience on an interactive writing canvas in Practice section (see figure 6 for an example). Along with hand-on learning practice, practice and corresponding English meanings, pictures of all eight Chinese characters were presented to help the participants recognize and understand the meaning of each character (see Appendix B for other seven characters presented in the Practice Section).

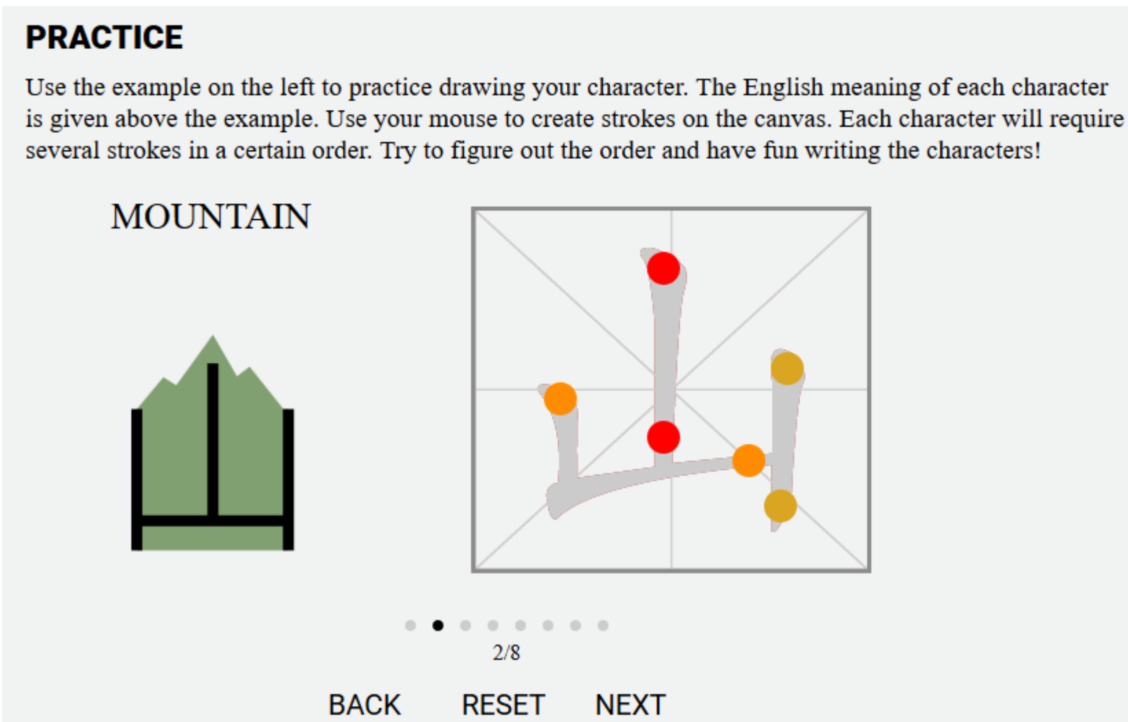
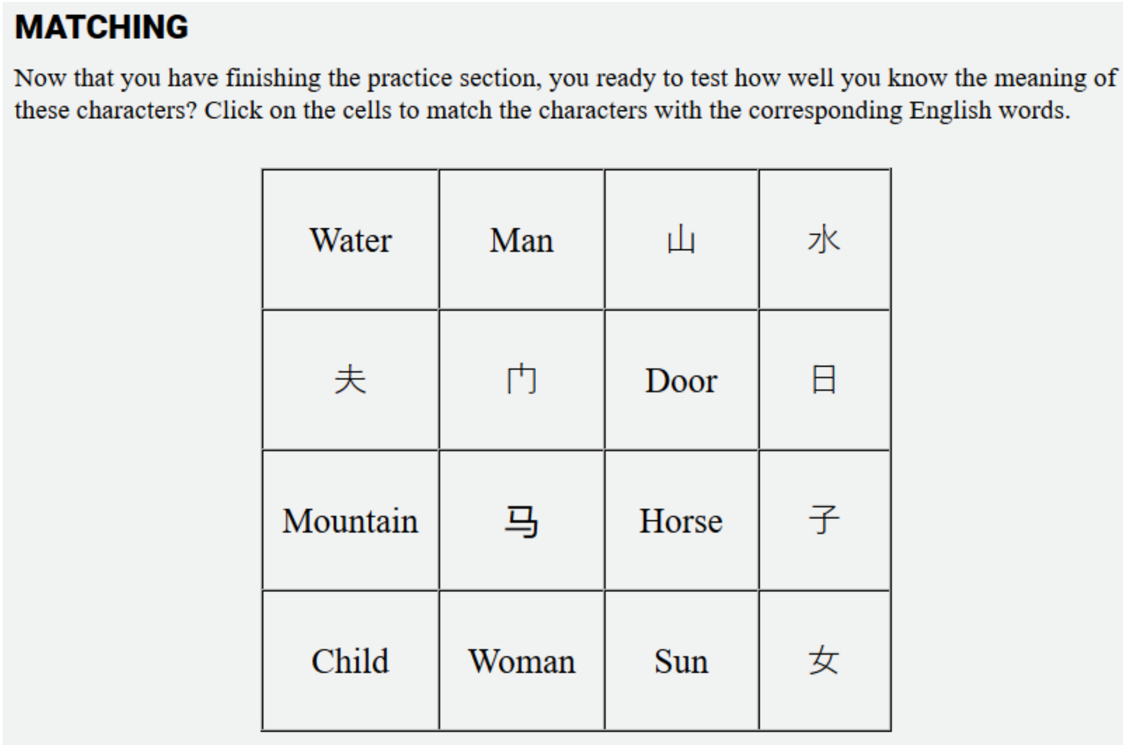


Figure 6. An Example of Interactive Writing Canvas in Practice Section

The knowledge checks included are a matching and a stroke order drag-and-drop activity. The matching activity required the participants to pair each Chinese character with its corresponding English meaning (See figure 7).



MATCHING

Now that you have finishing the practice section, you ready to test how well you know the meaning of these characters? Click on the cells to match the characters with the corresponding English words.

Water	Man	山	水
夫	门	Door	日
Mountain	马	Horse	子
Child	Woman	Sun	女

Figure 7. Matching Section

The Stroke Order knowledge check was a drag-and-drop activity that had the participants apply strokes to a box according to the correct stroke order for each Chinese character (see figure 8 for an example).

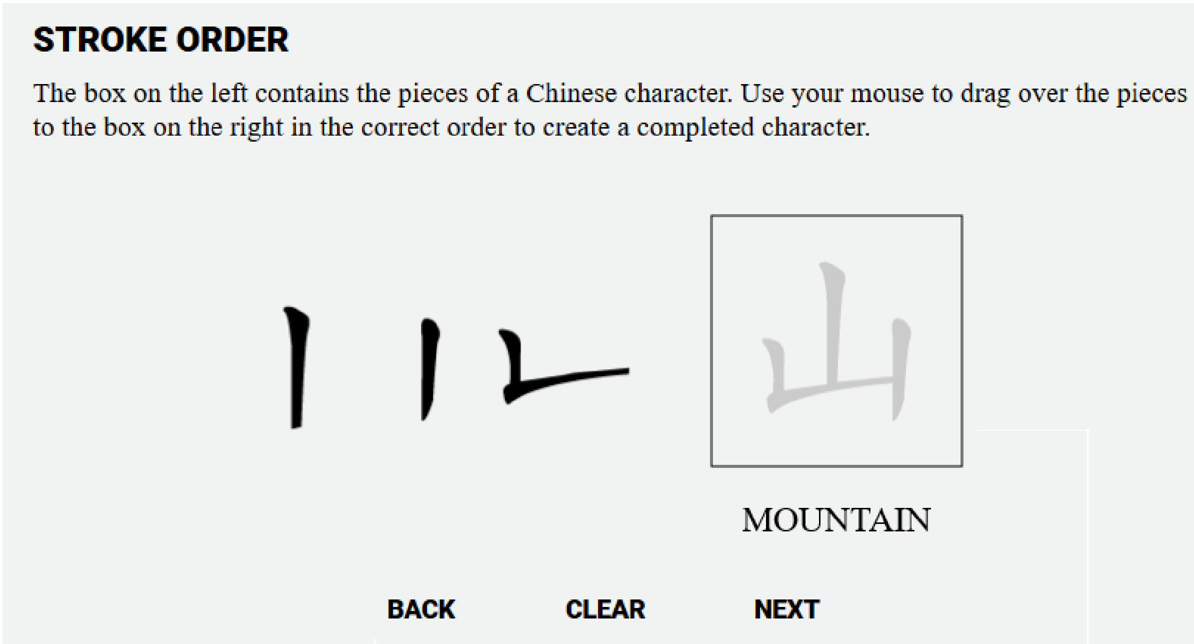


Figure 8. An Example of Stroke Order Assessment

According to the research design, the manipulated independent variable of this study was the version of the *Learning Chinese* program. Consequently, two versions of the *Leaning Chinese* program were created: the one with non-speech audio was for the experimental group, and the other without non-speech audio was for the control group.

3.3.2 The Version with Non-Speech Audio

The experimental group received a 30-minute instruction using the version of the *Learning Chinese* program that included non-speech audio. The foundation of non-speech audio design was in the learner-environment interaction. Therefore, the learner-environment interaction was analyzed in the Practice, Matching, and Stroke Order sections to understand what actions the participants might perform and what feedback should be sent to the participants to respond to each possible action. An interaction analysis was performed to examine where to implement non-speech audio.

The Practice section was the most essential section. Learners were expected to gain knowledge of the fundamental strokes and the correct order of eight entry-level Chinese characters by interacting with the designated canvas in this section. Each character required several strokes in a certain order and learners used the mouse to create these strokes on the canvas. An example of this was 山 (Mountain). The 山 character has three strokes. First, the center vertical stroke is drawn from the top downwards (1, 2). Then, the left wing is drawn from the top downwards before turning to continue from left to right (3,4). Finally, the right wing is drawn from the top downwards (5,6) (see figure 9).

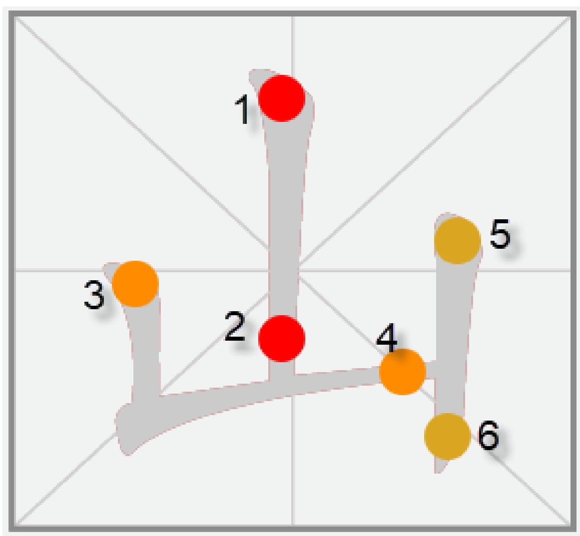


Figure 9. An Example of the Practice Section

During learner-environment interaction, three types of non-speech audio were embedded to modify and guide participants' writing through dynamic interactions with the writing canvas. One type of non-speech audio was used to confirm the correctness of the participants' movement, a second type on non-speech audio signified incorrect movements, and a third type indicated when the writing of the character has been completed. Initially, pre-recorded non-speech audio files were implemented in the *Learning Chinese* program. As audio files were generally larger than other media types, they tended to make the website heavier and slower to load. The amount of time spent waiting for these files to load had a negative impact on the learning experience because the auditory information was delayed. The Web audio application programming interface (Web Audio API) synthesizes audio using JavaScript. Such technique allows generating audio from scratch without increase any file size. What's more, Web Audio API has a low latency timing model that enables a fast auditory response to learner actions (Smus, 2013). Therefore, web audio API technique was applied to generate all types of non-speech audio in this study. For example, in the Practice section, the non-speech audio used to indicate an incorrect movement in the Practice section was coded as *playSound ('sawtooth', 60.13, 0, 0.12, false)*. The non-speech audio programmed in JavaScript eliminated the problems caused by large file sizes and avoided time spent waiting for files to load, and it helped to achieve the synchronization of visual, auditory, and tactile senses in the interaction process.

Figure 10 below describes the use of three types of non-speech audio when participants interact with the canvas to explore the stroke orders of the 山 character (mountain). Non-speech audio S1 confirmed the participants' movement so that they knew to continue to write next stroke, whereas non-speech audio S2 enabled the participants to know the negative results of their movement. Non-speech audio S3 informed the participants that they wrote the character in the correct stroke order.

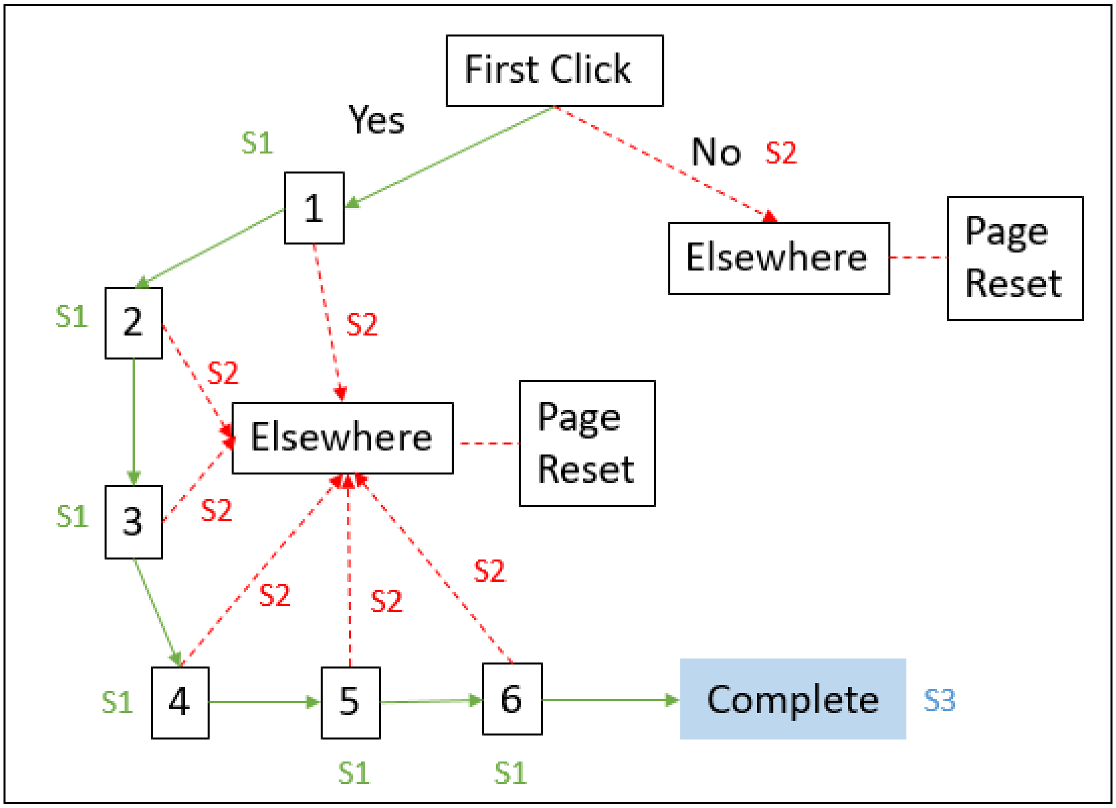


Figure 10. The Implementation of Non-Speech Audio in the Practice Section

The matching and Stroke Order sections were the knowledge check activities that enabled participants to verify their understanding of the Chinese characters. The results of the learners' actions could be categorized into three types: correct match (Drag), incorrect match (Drag), and match completed (Drag-and-Drop Completed).

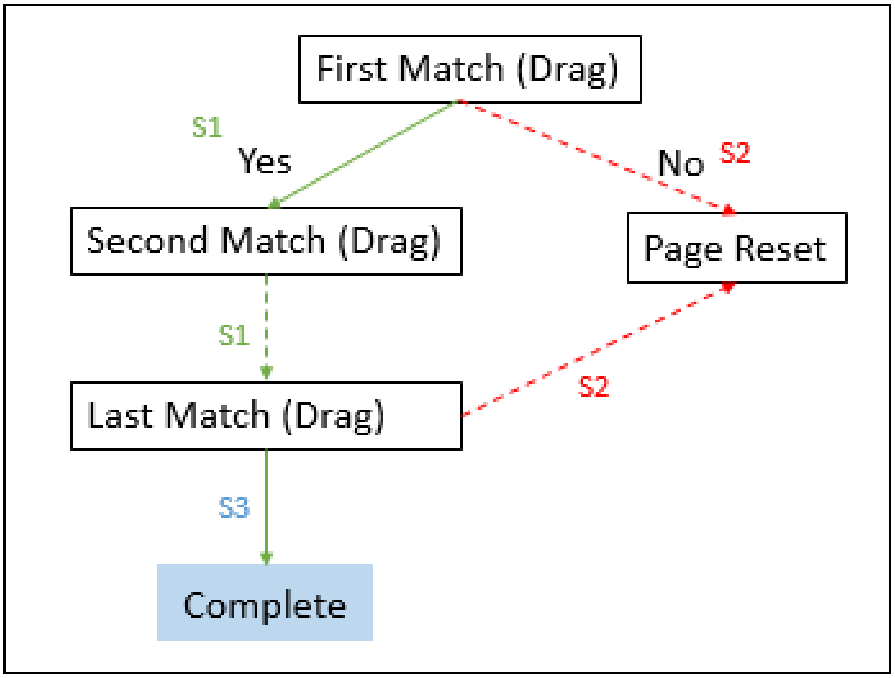


Figure 11. Non-Speech Audio in Knowledge Check Activities

Figure 11 portrays the use of non-speech audio in the knowledge check activities. Three types of non-speech audio were embedded to help the participants understand the results of their actions. Non-speech audio S1 confirmed the match (Drag) that the participants knew to continue to next pair (Stroke); non-speech audio S2 indicated the mismatch (incorrect Drag), and non-speech audio S3 implied the successful completion of the activity.

3.3.3 The Version without Non-Speech Audio

The control group received the same instructional content and knowledge activities except that it did not have any non-speech audio embedded. The equivalent visual representations (text and colors) were used to fill the non-speech audio void.

In the Practice section, the equivalent text information was provided when participants interacted with the canvas to explore the stroke orders. The text feedback was placed at the bottom of the canvas to modify and guide participant movement during the interaction.

Specifically, non-speech audio S1 was replaced by 'Correct! Continue!'; non-speech audio S2 was substituted by 'Incorrect! Start Over!', and 'You Got It! Next!' was used instead of non-speech audio S3 (see figure 12 for an example of text equivalent).

Figure 12. An Example of Text Equivalent

The knowledge check activities used colors to indicate the participants' success and failure movements. Matched cells (Stroke) would turn green and mismatched cells (Stroke) would turn red (see figure 13 for an example of the visual equivalent).

花	鸟	Sun	Man
人	Water	Child	子
Bird	水	女	Flower
舍	House	日	Woman

Figure 13. Matching Activity

3.4 Data Collection Procedure

The first step was to apply to Institute Review Board (IRB) to ensure this research complied with the research ethics protocols established by the Texas A&M University. Once permission was granted from the IRB, I commenced with data collection in the spring and summer of 2018. Due to convenient accessibility, participants were recruited from the Department of Educational Psychology at Texas A&M University. First, an invitation email was sent to the instructors of EPSY 433, INST 301, and EPSY 435 (see email in Appendix G) to introduce this research study as well as the benefits of participation. A sign-up Google spreadsheet was also included in the invitation email. Second, the instructors forwarded the invitation email to their students, and those who were interested in the study could provide their names and email addresses in the Google spreadsheet. As of July 2018, 140 students signed up to participate in this study. Third, the recruited participants were randomly assigned to the control group (A) or experimental group (B) using a random number generator in Excel. Each group received step-by-step instructions directly from me on how to take the *Learning Chinese* program (see instructions in Appendix H). Fourth, the participants completed the 30-minute *Learning Chinese* program at a location of their choice, using their own electronic devices. Before the participants were able to exit the *Learning Chinese* website, they were asked to complete a brief 5-minute online survey. While the participants were going through the website, web data were being collected at the backend. Therefore, two types of data were collected in this study: one was the data from the questionnaire entered by the participants in Qualtrics and the other was the web data recorded at the back end of the *Learning Chinese* program.

3.5 Instrumentation

The instruments for this study are the User Experience Questionnaire (UEQ), the Situational Motivation Scale (SIMS), and web data.

3.5.1 User Experience Questionnaire

The User Experience Questionnaire is constructed based on a theoretical framework of Human-Computer Interaction. This research framework focuses on the user's subjective perception of product features when interacting with a product and the immediate impact of such subjective perception on the user him/herself (Hassenzahl, 2007; Laugwitz & Schrepp, 2008). This research framework suggests ergonomic quality and hedonic quality as key dimensions of user experience (Hassenzahl, 2007). Ergonomic quality measures perceived usefulness and sheds lights on the users' perception of the product's ability to help them to complete a task with efficiency and effectiveness. On the other hand, hedonic quality measures the non-task-oriented aspects of a product such as the originality of the product design.

The UEQ measures user – interface interaction experience in online environments that allows the users to express feelings, impressions, and attitudes that arise when interacting with the product in a simple and immediate way. As outlined in figure 15, the questionnaire contains 6 factors: Attractiveness, Efficiency, Perspicuity, Dependability, Stimulation, and Novelty.

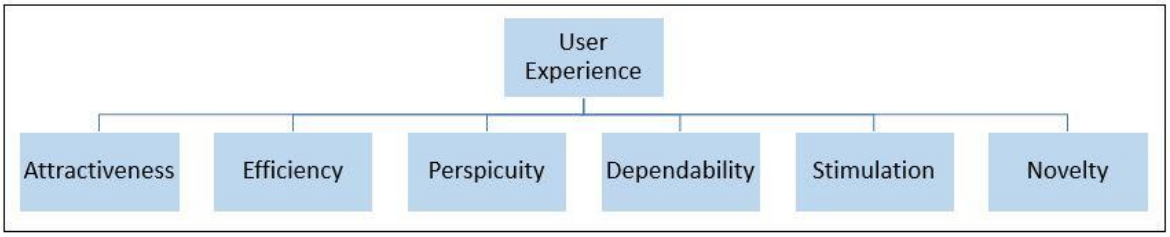


Figure 14. Six scales of User Experience (Schrepp et al., 2017)

The factor of Attractiveness represents a purely emotional reaction that doesn't explain any reason for the acceptance or rejection of the product. For example, items such as good versus bad, and pleasant versus unpleasant. However, with Efficiency, Perspicuity, and Dependability factors, the focus is on the task-oriented aspect of the interaction experience. Specifically, these three factors examine the ways in which the product enables the users to complete the task with efficiency and effectiveness. For example, items such as easy to learn versus difficult to learn (perspicuity), fast versus slow (efficiency), and unpredictable versus predictable (dependability). Finally, the Stimulation and Novelty factors measure the non-task oriented aspects of the interaction experience such as the originality of the design. For example, items such as dull versus create (novelty), boring versus exciting (stimulation).

The final questionnaire contains six factors together with 26 items, using a 7-point rating scale. These six factors are extracted and analyzed from an initial 229 items using a factor analysis. The selected items have high loadings on the respective factor. The order and the polarity of the final 26 items are randomized in the questionnaire. An example of the items in the UEQ is:

annoying	○	○	○	○	○	○	○	enjoyable

The internal consistency reliability and validity of the UEQ was investigated in 11 usability tests with a total number of 144 participants and in an online survey with 722 participants. As reported in the User Experience Questionnaire Handbook (Schrepp et al., 2017), the results of the usability tests and the online survey indicated a sufficiently high scale internal consistency reliability. In addition, the survey development team conducted several studies where high positive correlations were observed, to validate the construct validity of the UEQ (Laugwitz et.al, 2008).

3.5.2 The Situational Motivation Scale

The Self-determination theory (SDT) has been one of the most influential theories used to investigate types of motivation that underlie human behavior (Guay, Vallerand, & Blanchard, 2000; Hartnett, George, & Dron, 2011). This theory conceptualizes a self-determination continuum from high to low levels and report on factors that may affect the quality of behavior such as amotivation, external regulation, identified regulation, and intrinsic motivation (see figure 16).

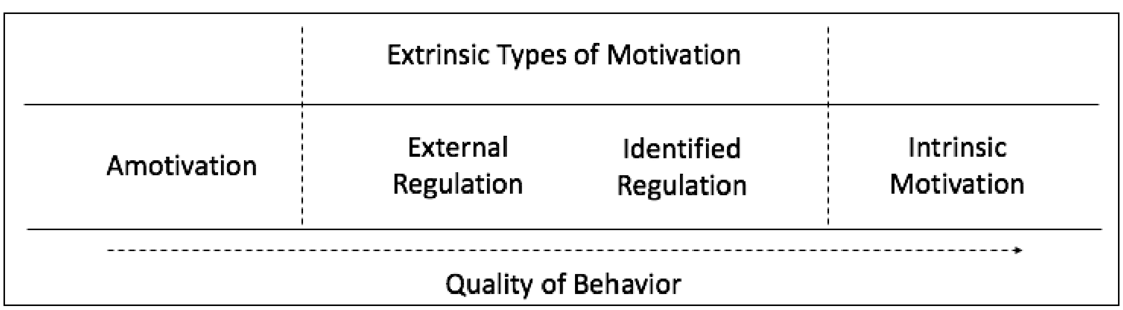


Figure 15. The SDT Model of Motivation

Intrinsic motivation refers to the pleasure and satisfaction originated while learners perform the activities (Deci, 1971). In contrast, extrinsic motivation relates to the behaviors initiated for outside of the activity itself. According to the level of self-determination, extrinsic motivation can be categorized into identified regulation and external regulation. Identified regulation is considered to be extrinsic motivation, but it is internal in the sense that learners are able to perceive relevance and worthwhile of the learning activities designed by the instructors (Brophy, 2010; Hartnett, George, & Dron, 2011; Ryan & Deci, 2000). External regulation is a type of extrinsic motivation that occurs when rewards or punishment are involved (Hartnett, George, & Dron, 2011). Finally, amotivation pertains to the least motivated behaviors that learners experience and is similar as learned helplessness (Abramson, Seligman, & Teasdale, 1978; Guay, Vallerand, & Blanchard, 2000).

Many research studies have concluded that the self-determination theory is a useful analytic tool for the study of the complexity of motivation and could be used to explore learners' reasons for engagement in online learning environments (Chen & Jang, 2010; Hartnett, 2010; Hartnett, George, & Dron, 2011; Martens et al., 2004; Rovai et al., 2007; Shroff et al., 2007; Shroff et al., 2008; Xie et al., 2006). The Situational Motivation Scale is constructed based on the self-determination theory. This self-reported survey includes four factors with 16 items, using a 7-point rating scale. A factor analysis was performed to investigate the factor structure for the SIMS (see Loadings of the items on the factors). The examples of the items in the SIMS are listed in table 2.

The internal consistency of the items was measured with Cronbach's alpha and the results indicated a sufficiently high internal consistency reliability (Guay, Vallerand, & Blanchard, 2000). For construct validity, strong positive correlations were observed in three correlational analyses conducted by Guay et al (2000). Since this study focused on the impact of non-speech audio impacted on learners' motivation, the factors of intrinsic motivation and identified regulation from the SIMS were used in this study.

Table 2. An Example of the Items

Why are you currently engaged in this activity?							
1. Because this activity is fun	1	2	3	4	5	6	7
2. Because I think that this activity is good for me	1	2	3	4	5	6	7
3. Because I am supposed to do it	1	2	3	4	5	6	7

3.5.3 Web Data

The amount of time spent on the webpage has been considered as one of the key indicators of the level of engagement or involvement that a user has with a website (Liu, White, & Dumais, 2010). The information recorded via website is seen as a type of objective data that is useful to understand online behavior. Therefore, a WebStats scripts were embedded in the *Learning Chinese* program to track learners' online behavior.

user name	visit time	keep time(seconds)	status
[redacted]	2018-06-11 21:36:54	44	completed
[redacted]	2018-06-13 21:33:06	39	completed
[redacted]	2018-06-15 00:53:32	32	completed
[redacted]	2018-06-16 14:36:17	16	left
[redacted]	2018-06-16 14:38:35	3	left
[redacted]	2018-06-16 21:59:20	29	completed
[redacted]	2018-06-17 02:37:24	36	completed

Figure 16. An Example of Web Data Collected at the Backend

As figure 17 shows, when a participant visits the *Learning Chinese* program, time spent on each page, the status of each learning task, and visit time are automatically recorded at the backend of the *Learning Chinese* program.

Time spent on Matching and Stroke Order activities were used to understand whether the use of non-speech audio helps to improve the efficiency of learner performance. Time spent on the Practice section was used to understand how the use of non-speech audio helps to engage learners in learning the content.

3.6 Data Analysis

The data collected from the *Learning Chinese* program back end were cleaned and the responses to the questionnaire were converted to numerical values using Microsoft Excel. And then, the data were imported for analysis into SPSS version 22. The initial data analysis phase was to assess the reliability of the UEQ and the SIMS. Cronbach's alpha of scales were obtained from the statistical tests and were compared with the results in the published literature.

In the main data analysis phase, the aim was to answer the research questions and to provide insights on the effect of non-speech audio as used in an online learning environment.

Specifically, the analysis was done to determine whether the addition of non-speech audio in the interaction process enhanced the learner's online experience (UEQ), stimulated motivation (SIMS), and encouraged learners to engage in learning (Time on Practice), and thereby led to better learning performance (Time on Matching and Stroke Order). The analysis of the main effects of the independent variable (version of the program) on the dependent variables (learner experience, learner motivation, and time spent on the program) was guided by the following research hypotheses:

- There is no statistically significant difference on the UEQ between the average score of the undergraduate students who used the non-speech audio version of *Learning Chinese* program and the average score of the undergraduate students who used the version of *Learning Chinese* program without non-speech audio.
- There is no statistically significant difference on the SIMS between the average score of the undergraduate students who used the non-speech audio version of *Learning Chinese* program and the average score of the undergraduate students who used the version of *Learning Chinese* program without non-speech audio.
- There is no statistically significant difference between the amount of average time spent on the program by the undergraduate students who used the non-speech audio version of *Learning Chinese* program and the amount of average time spent on the program by the undergraduate students who used the version of *Learning Chinese* program without non-speech audio.

The analysis of the interaction effects of the independent variable (version of the program, self-perceived learning styles, attitude toward online learning) on the dependent variables (learner experience, learner motivation, and time spent on the program) was guided by the following research hypotheses:

- There is no statistically significant difference in undergraduate students' average score on UEQ, SIMS, and Time spent on the program based on the version of the *Learning Chinese* program student participated in and their self-perceived learning styles.
- There is no statistically significant difference in undergraduate students' average score on UEQ, SIMS, and Time spent on the program based on the version of the *Learning Chinese* program student participated in and their attitude towards online learning.
- There is no statistically significant difference in undergraduate students' average score on UEQ, SIMS, and Time spent on the program based on the version of the *Learning Chinese* program student participated in, self-perceived learning styles, and their attitude toward online learning.

A multivariate analysis of variance (MANOVA) was performed to answer and interpret the research hypotheses listed as above. MANOVA is a way to test the hypotheses that one or more independent variables has an effect on two or more dependent variables (O'brien and Kaiser, 1985). A MANOVA enables us to better discover which factor is truly important when measuring several dependent variables in a single experiment. Moreover, a MANOVA can provide protection against Type I errors that might occur if multiple univariate analysis were conducted independently. When using a MANOVA, the dependent variables need to be significantly related. Additionally, the observed covariance matrices of the dependent variables should be equal across groups. Therefore, Bartlett's Test of Sphericity[3] was used to examine whether the dependent variables are related, and Box's *M* was used to test the homogeneity of covariance matrices in this study.

In the main analysis phase, Wilk's lambda (λ) was used to test the statistical significance between the means of identified groups on a combination of dependent variables. If the overall multivariate test is significant, it can be safely concluded that the respective effect (i.e., the effect of non-speech audio) is significant. In addition to λ, partial eta squared (η^2) was used to examine the magnitude of the effect of non-speech audio on the dependent variables. After obtaining an overall significance, a further analysis was conducted to identify how the specific dependent variables contributed to the significant overall effect. Furthermore, the values of p were obtained to identify a statistical interaction effect among the independent variables (version of the program, self-perceived learning styles, attitude toward online learning).

CHAPTER IV

RESULTS

This chapter outlines the results of the data analyses conducted to address the research questions and hypotheses. It begins with a report of Cronbach's alpha of the scales to validate the internal consistency of the UEQ and the SIMS, and reflects on the results of a 2 x 3 MANOVA are presented to answer the research questions and hypotheses.

4.1 A Report of Cronbach's Alpha

Cronbach's Alpha coefficient was calculated for each subscale of both the UEQ and SIMS, using Scale – Reliability function in SPSS.

For the UEQ, Cronbach's alphas for six subscales were .90 (attractiveness), .85 (perspicuity), .79 (efficiency), .77 (dependability), .91 (stimulation), .88 (novelty), respectively. The alphas for all six subscales were higher than .75, suggesting that the subscales had a high level of inter-item reliability. Further analysis found that deleting any of the items would not have significantly increased alpha.

For the SIMS, Cronbach's alphas for two subscales used in this study were .92 (intrinsic) and .86 (identified regulation), respectively. The alphas for both subscales were higher than .85, which indicated that the subscales had a high level of inter-item reliability. The results of Cronbach's Alpha if Item Deleted displayed that deleting any of the items would have decreased the alpha level.

4.2 Descriptive Statistics

There were 140 participants randomly assigned to the experimental group (n=70) and control group (n=70) in this study. On average, the overall average of scores on learner motivation from the treatment group (M =46.34, SD = 7.19) was higher than those of the control group (M = 39.69, SD = 9.71). In addition, when comparing mean scores among the participants with the same learning styles between the two groups, scores from the treatment group were all higher than those of the control group (see Table 3).

Table 3. Descriptive Statistics of Learner Motivation

	Control			Treatment		
	N	M	SD	N	M	SD
Learner Motivation						
Visual	28	40.07	9.10	31	45.32	5.22
Audio	7	32.43	13.59	12	47.92	7.50
Kinesthetic	16	41.50	8.31	14	47.29	7.62
Multi Selection	19	40.26	9.70	13	46.31	6.90
Total	70	39.69	9.71	70	46.34	7.19

Similarly, the overall average of scores on online learning experience from the treatment group (M = 156.57, SD = 20.59) was higher than those of the control group (M = 147.76, SD = 24.39). Moreover, for the participants who perceived themselves as visual learners, auditory learners, and kinesthetic learners, mean scores on online learning experience from the treatment group were all higher than those of the control group (see Table 4). For the participants who perceived themselves to have multiple learning styles, the treatment group had an average of 161.92, which was higher than the average of 157.26 from the control group. However, the standard deviation of the mean score in the treatment group (SD = 22.33) was slightly larger than that of the control group (SD = 20.59).

Table 4. Descriptive Statistics of Online Learning Experience

	Control			Treatment		
	N	M	SD	N	M	SD
Online Learning Experience						
Visual	28	140.71	22.42	31	155.68	20.70
Audio	7	133.86	27.84	12	149.42	21.84
Kinesthetic	16	154.88	25.71	14	159.71	17.59
Multi Selection	19	157.26	20.59	13	161.92	22.33
Total	70	147.76	24.39	70	156.57	20.59

As displayed in table 5, the participants in the treatment group had an average time of 636.11 seconds spent on the *Learning Chinese* program (SD = 302.65), and those of participants in the control group had an average time of 578.47 seconds (SD = 191.57). Although the participants from the treatment group spent more time on the program than those of in the control group, the standard deviation of the mean score (SD = 302.65) in the treatment group was larger than the standard deviation of mean score in the control group (SD = 191.57). When comparing time spent on the program among the participants with the same learning styles between two groups, the participants from the treatment group spent more time on the program than those of the control group, yet the standard deviations of mean scores were larger in the treatment group than those of in the control group.

Table 5. Descriptive Statistics of Time

	Control			Treatment		
	N	M	SD	N	M	SD
Time Spent on the Program						
Visual	28	587.82	209.76	31	652.87	281.08
Audio	7	569.29	162.81	12	658.00	354.48
Kinesthetic	16	556.50	131.57	14	606.07	257.01
Multi Selection	19	586.58	202.79	13	608.31	373.28
Total	70	578.47	191.57	70	636.11	302.65

4.3 Checking the Assumptions of MANOVA

Box's Test of Equality of Covariance Matrices checked the assumption of homogeneity of covariance across the groups using p <.05 as a criterion. The results exhibited that Box's M (.418) was not significant, p (.51) > (.05), indicating that there were no significant differences between the covariance matrices. Therefore, the assumption was not violated, and Wilk's Lambda was appropriately used to determine the significance level.

Bartlett's Test of Sphericity[3] measured the strength of the relationship between the dependent variables using p <.05 as a criterion. Based on MANOVA results, Bartlett's Test of Sphericity[3] was significant, p (.000) < (.05), indicating that the correlation matrix was not an identity matrix. In other words, a sufficient correlation was found among the dependent variables. Therefore, the condition of Sphericity was not violated and the MANOVA performed in this study would not risk increasing the likelihood of a Type I error.

4.3.1 Results of Main Effect

For the purpose of this study, the version effects are of interest because they tell us whether or not the use of non-speech audio has an impact on three dimensions of learner's online experience (UEQ), learner motivation (SIMS), and time spent on the program. According to MANOVA results reported in SPSS (see Table 6), using an alpha level of .05, the version effects suggested that this test was significant, Wilk's Λ = .841, F (3, 121) = 7.619, multivariate η^2 = .159. Power to detect the effect was .986. This result indicated that there were significant differences between the two versions of the program on a linear combination of the three dependent variables. The multivariate η^2 = .159 indicated that approximately 15% of the multivariate variance of the dependent variables was associated with the factor of version of the program. Such effect size was considered a relatively large effect size (Watson, 2016)

Table 6. Multivariate Test

Effect	Value	F	Df	Sig.	Partial Eta Squared	Observed Power[d]
Versions_Wilk's Lambda	.841	7.619	3	.000	.159	.986

Given the significance of multivariate effect, this effect needed to be broken down to determine the individual dependent variables with separate ANOVA tests. Prior to examining the results of separate ANOVA tests, it was important to first test the assumption of homogeneity of variance for all the dependent variables. In other words, Levene's Test needed to be non-significant for each of the dependent variables if homogeneity of variance had been met. The results displayed that this assumption was met for all three dependent variables, which confirmed the reliability of the univariate tests to follow (see Appendix F figure 27 Levene's Test results for all three dependent variables).

The univariate main effect was significant for learner motivation, $F = 14.619$, $\eta^2 = .106$, $p = .000$, Power to detect the effect was .967. The $\eta^2 = .106$ indicated that approximately 10.6% of the variance of the learner motivation was associated with the factor of the version of the program (see Table 7). As Watson (2016) suggested, a partial eta squared of .106 represented a medium effect size. Likewise, the univariate main effect was also significant for learner online experience, $F = 7.994$, $\eta^2 = 0.61$, $p =.005$, Power to detect the effect was .801. The $\eta^2 = .061$ indicated that approximately 6.1% of the variance of the learner online experience was associated with the factor of the version of the program. However, the univariate main effect for time spent on the program (web) was insignificant, $F =.852$, and $p = .358$.

Table 7. Tests of Between-Subject Effects

Source	Dependent Variables	df	Mean Square	*F*	Sig	Partial Eta Squared	Observed Power[d]
Versions	Online Experience	1	3979.293	7.994	.005	.061	.801
	Learner Motivation	1	1129.547	14.619	.000	.106	.967
	Time Spent on Program	1	54860.326	.852	.358	.007	.150

4.3.2 Results of Interaction Effect

The analysis of the interaction effects was guided by three null hypotheses which assumed no statistical interaction among the independent variables of the version of the program, learner styles, and learner' attitude toward online learning.

Table 8. Interaction Effect

Effect	Value	F	Df	Sig.	Partial Eta Squared	Observed Power[d]
Versions*Learning Styles	.950	.699	9	.709	.017	.348
Versions*Attitude	.979	.857[b]	3	.465	.021	.232
Effect	Value	F	Df	Sig.	Partial Eta Squared	Observed Power[d]
Versions*Learning Styles*Attitude	.984	.319[b]	6	.926	.008	.139

The results showed Wilk's ʌ for the interaction of the version and learner styles was .950 and *p* = .709; Wilk's ʌ for the interaction of the version and attitude was .979 and *p* = .465; and Wilk's ʌ for the interaction of the version, learner styles, and attitude was .984 and *p* = 926 (see Table 8). According to these results, neither interaction effect was statistically significant using alpha level at .05. Specifically, the effect of the independent variable (the version of the program) on the dependent variables (the UEQ and SIMS) was not dependent on the level of the other independent variables (learner styles and attitude). Therefore, we failed to reject all three null hypotheses about the interaction effect.

4.4 Summary

This chapter examined the results of data analyses of the study. A summary of descriptive statistics was presented first to understand the difference of the mean scores on all three dependent variables (i.e., UEQ, SIMS, and time) between the control group and the treatment group. Overall, the mean scores of all three dependent variables were higher than those of in the control group, except that the standard deviations of the mean scores for time spent on the program in the treatment group were higher than those of in the control group.

The analysis of the main effects was guided by three null hypotheses which assumed no statistically significant difference between the mean scores of the control group and the mean scores of the treatment group on the UEQ, SIMS, and the Time on the program respectively. According to the results of a 2 x 3 MANOVA, the version of the program had a significant effect on UEQ and SIMS. The null hypotheses on UEQ and SIMS were rejected. However, the version of the program has no significant effect on time spent on the program, and we failed to reject this null hypothesis. For the interaction effects, neither of the three null hypotheses was statistically significant, as a result, we failed to reject all three null hypotheses about interaction effects among the independent variables.

CHAPTER V

CONCLUSION

Non-speech auditory messages can communicate information to learners without interrupting the learning process. Yet, researchers and practitioners rarely consider using non-speech audio in the learning environment. As revealed in the systematic literature review in this study, in the past decades, only a few studies have investigated the effect of non-speech audio on learning, and mainly, as auditory stimuli embedded in less complex, linear learning environments. Given the development of interactive learning environments, the potential impact of non-speech audio involves more than merely attracting learners' attention as a stimulus. Therefore, this dissertation aimed to explore the effects of non-speech audio used in interactive online learning environments.

The study is guided by two research questions which are investigated through six hypotheses. An instructional program was designed to evaluate the effect of non-speech audio when implemented in the interaction process. An experimental design was followed for data collection from 140 Texas A&M University undergraduate students. The treatment group used the version of the program with non-speech audio embedded in the interaction process, whereas the control group used the version with equivalent visual information to fill the void of the sound in the interaction process. During the experiment, questionnaires were administered, and website data was recorded for data analysis. All these details were described in the previous chapters.

In this chapter, a summary of the major findings for each hypothesis is presented, along with a discussion of the implications of these findings. Moreover, suggestions are made for future research on the use of non-speech audio for learning.

5.1 A Summary of the Major Findings

In this study, the version of the program was the main predictor of the effect of non-speech audio across all three behaviors as discussed in preceding chapters: online interactive experience, learner motivation, and time spent on the program. Moreover, interaction effects were investigated to understand whether the effect of non-speech audio differs depending on the level of learner styles and attitudes towards online learning.

5.1.1 Non-Speech Audio and Online Interactive Experience

The findings indicated that the addition of non-speech audio in the interaction process has a positive effect on learners' online experience. In particular, non-speech audio impacts the way learners experience their interaction with the instructional program. In a virtual learning environment, learner experience is built through interaction with the elements of the environment. Learner's perception and feelings towards this interaction originate primarily from their impressions of the product's design and use.

The UEQ measures learners' impression on the design of the program. The findings suggest that the addition of non-speech audio in the interaction process have a positive effect on the originality of the design. In this study, learners gained knowledge of the fundamental strokes and the correct stroke orders of Chinese characters by interacting with the designated canvas. Non-speech audio was embedded in this dynamic interaction process to confirm and modify learners' actions. In addition, programmed non-speech audio enabled the synchronization of visual, auditory, and tactile senses in the interaction process. Participants who used the version of the program with non-speech audio were able to communicate with the instructional program while engaging in a multi-sensory learning experience. This design simulated ways in which a person would interact with everyday tasks in a real-life environment, which differed from many online learning programs (e.g., the version used for the control group) that information was delivered through visual representations. Consequently, the use of non-speech audio in the interaction process gave the learners a sense of the innovative design of the instructional program.

Additionally, the UEQ also examines learner's subjective perception of the use of the program. Here findings suggested that the use of non-speech audio has positively influenced learners' impression on the use of the program. For example, the use of non-speech audio seemed to enable the learners to complete learning activities with greater efficiency and effectiveness. Participants who used the version of the program with non-speech audio, developed their understanding of the stroke orders through both visual and auditory channels. While performing the main drawing task, immediate feedback was given through the auditory channel. By doing so, participants were able to maintain their attention on the drawing activity. On the contrary, participants who used the version of the program without non-speech audio had to direct their visual attention to the feedback area constantly. Even though the text feedback was presented close to the drawing canvas in a succinct way, a frequent attention switching could disrupt information encoding and lead to ineffective learning (Hogan, Kelly, &Craik, 2006).

As discussed in the literature review, the effects of non-speech audio were primarily examined in linear learning environments where learner-environment interaction was limited or not required. Consequently, studies did not yet investigate ways in which the use of non-speech audio could impact learner experience in an interactive virtual environment. Therefore, the findings which suggest that the addition of non-speech audio in the interaction process has a positive effect on learners' online experience, is novel to the field of education.

5.1.2 Non-Speech Audio and Learner Motivation

In this study, learner motivation refers to internal pleasure and satisfaction experienced while participants work through learning activities. A significant difference in motivation was detected in participants of the treatment group and the control group. This finding indicates that the addition of non-speech audio in the interaction process has a significant effect on learner motivation. The only difference between the two versions of the program was the design of response to the learner's action. One version of the program used non-speech audio to confirm or redirect the learners' action. While learners drew their stroke, a continuous ringing sound was played to confirm correct strokes. A low heavy "Durr" sound was used to redirect incorrect strokes, and a bell sound was used to confirm the successful completion of a Chinese character. The other version of the program used equivalent text information to guide learner's behavior during the drawing practice. In this respect, non-speech audio seems to be a more effective way to trigger the emotions such as excitement and accomplishment that motivated learners to explore the drawing of the Chinese characters in this study.

Bishop (2000) observed a similar effect and suggested that the use of non-speech audio in computer-based lessons helped unmotivated participants to actively explore the learning materials and remain interested after the completion of the lesson. Similarly, Königschulte (2015) also found non-speech audio to have a positive influence on learner motivation and engagement in the learning process. Certainly, differences in the design and selection of non-speech audio existed between this study and two research studies. For example, these researchers designed and investigated the use of non-speech audio as auditory stimuli in computer-based learning environments with limited interaction.

5.1.3 Other Findings

No interaction effect was detected among the versions of the program, learner styles, and learner attitude in this study. Learner styles and their attitudes are considered influential factors on learners' experience in online courses (Cereijo, 2006; Song et al., 2004). The results of this study indicate that the effect of non-speech audio would not be significantly altered by learner style differences or by attitudes towards online learning. Simply stated, when non-speech audio is added to the interaction process, positive effects on learner online experience and learner motivation apply to learners with different learner styles and attitudes toward online learning.

Unlike the above-mentioned positive effects, the results of this study reveal that the amount of time spent on the instructional program did not differ significantly between the two groups. The amount of time spent on a product is considered to be an objective indicator of the level of engagement that learners have with the product (Liu, White & Dumais, 2010). Similarly, in the current study, results from the participants' self-reported questionnaire suggest that non-speech audio could be an effective motivator to learners as they engage in the learning process. However, data generated from the program's back end found that participants who used the program with non-speech audio did not spend more time than those who used the program without non-speech audio. A possible explanation could be that only eight entry-level Chinese characters were used in this study. Based on our pilot testing, the average amount of time for a non-native speaker to discover the stroke order of one Chinese character was around 15 seconds. As a result, time required to complete all eight Chinese characters may not have been enough to lead to a conclusive understanding.

5.2 Implications of the Study

This study is the first attempt to investigate the effects of non-speech audio in the interactive learning environments. The findings of this study suggest two main implications for researchers and practitioners who are interested in the design of effective interactive virtual learning environments. The task of learning to write Chinese characters was selected as the subject for the research purpose of this study, but the design and development of non-speech audio can be applied to the development of interactive learning system for use in a wide range of environments such as academia or corporate learning and development.

The first implication affects the impact and value added associated with the use of non-speech audio in the design of virtual learning environments. Non-speech audio, as a basic source of information, can be used to communicate important details and to reinforce information given to the learner in various learning settings. According to the findings of this study, a systematic application of non-speech audio in the interaction process is beneficial for learning because it enhances learners' online experience and motivates them to engage in the learning process. In addition, the use of non-speech audio enables the learner to process more information because it does not add to the visual information load in the way that the text-based feedback back during the dynamic interactions. Today's emerging technologies, such as augmented reality and virtual reality, allow for the development of interactive learning environments through which learners can interact with their surroundings in similar ways as they would in real life. For this reason, the implementation of non-speech audio in the interaction makes the learning experience more authentic and engaging, which would thus also facilitate retention and knowledge transfer. Hence, it is suggested that designers should consider integrating non-speech audio into interactive learning environments in ways that will use non-speech audio as the optimal format to deliver essential messages.

Moreover, when designing a linear learning environment, non-speech audio could play a significant role in the facilitation of effective learning. According to existing literature, the use of non-speech auditory stimuli could direct learners' attention to specific information, emphasize key concepts, and structure learning information, particularly when learners need to focus attention on learning in a linear learning environment.

The second implication relates to the theoretical framework which guided the design of non-speech audio in this study. The lack of effective design principles is an impediment to the systematic and effective implementation of non-speech audio in different learning settings. Design by intuition is one of the main approaches used by many designers and researchers in the field of education (Mann, 2008). However, the implementation of non-speech audio in learning environment by sheer intuition could produce inconsistent outcomes because decisions are based upon personal assumptions rather than on proven experience. In this study, the design of non-speech audio was guided from the perspective of learner interaction and non-speech audio was embedded into an interactive learning environment to allow for more meaningful and effective learning. Consequently, the design of non-speech audio was a part of the interaction analysis. The following questions were used to guide the design of non-speech audio in the interaction analysis in this study:

- What is the goal of an interaction?
- What are the possible actions or reactions that might be expected from the learners when they are presented a specific type of information?
- What are the possible roles that non-speech audio can play in a particular interaction?

- Is non-speech audio the best format to deliver the information?
- How will the sound be developed and implemented in the selected program?

Thus, it is suggested that designers consider the above guidelines when planning to integrate non-speech audio into an interactive learning system.

5.3 Limitation of the Study

Although the research has reached its aims, there were some unavoidable limitations.

First, the convenience sampling method was used to recruit the sample participants, limiting the generalizability of the results to a different population.

Second, the Chinese characters selected for the study was entry-level. Similarly, the drag and drop activity was used to examine whether the participants could write the selected characters in the correct orders. The drag-and-drop activity was not the best strategy to measure a deep understanding of learning efficiency and effectiveness. Initially, 15 Chinese characters were selected from a story, and the final assessment was to allow the participants to write on a blank canvas. This lack of complexity on content and assessment can limit the ability to determine true group differences in learning efficiency and effectiveness.

Third, the use of computers as delivery tool was not optimal for the topic used in this study. The decision to use computers was made based on cost and conveniences. Since the focus of this study was for the participants to learn how to write Chinese characters, the ideal delivery tool would have been touch screen devices that would have allowed the participants to draw the Chinese characters with their fingers instead of a computer mouse.

Forth, web data was insufficient to explain learners' online behavior patterns. Due to the limited time and resources, web data was limited to the amount of time spent on each page. The amount of time didn't provide any details of behavior patterns. For example, the navigational pattern tells learners' interest on specific topic; the amount of hits explains how engaged learners are on specific items. If web data were able to explain the details, we could understand the effects of non-speech audio in a more depth and objective way.

5.4 Recommendations for the Future Research

The goal of this study was to investigate the effects of non-speech audio in an interactive learning environment. Based on the results of this study, the following recommendations for future research include:

- An implementation of the complete conceptual model of non-speech audio design. This study focused on the second half part of the conceptual model that examined how to use the non-speech audio to respond to learners' actions in the interaction process. According to conceptual model, non-speech audio could be used as a stimulus to direct learners' attention to the specific learning tasks. In addition, non-speech audio was effective in stimulating learners' attention in linear learning environment according to the existing literature. Future research should investigate the effect of non-speech audio in an interactive learning environment in which the complete conceptual model of non-speech audio design can be implemented.
- More comparison groups to discover the most effective design. The researchers might consider having three groups: visual only, audio only, and visual + audio. In particular, when using non-text visual representation to convey the information, whether the effect of non-speech audio observed in this study still applies.
- The investigation of a sound-enhanced learning environment. Future research could investigate the effect of sound-enhanced learning environments where different types of

sound are implemented such as environmental sound, speech sound, interaction-based non-speech sound, and so on. Such an environment entirely simulates how a person interacts with everyday tasks in real-life. The researchers can separate the types of sound and investigate which combination of sound would have the most positive effect on learning.

- The use of web data to evaluate the effect of non-speech audio in learning. Future research could apply more complex techniques to record learners' online behavior patterns that provide more objective data to understand the effect of non-speech audio.
- Tests that measure a deep understanding of non-speech audio needed to be conducted. As initial attempt to explore the use of non-speech audio, this study primarily focused on learner experience and motivation. Future research should investigate how the use of non-speech audio could influence factors such as learner's attention, short-term memory, and long-term memory in the interactive learning environments.
- Verify the effectiveness of design guidelines of non-speech audio proposed in this study. The systematic application of non-speech audio used in this study was guided from the perspective of interaction. This design should be applied to other scenarios and subjects to verify its effectiveness.

REFERENCES

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of abnormal psychology, 87*(1), 49.

Aleksy, M., Bronmark, J., & Mahate, S. (2016, March). Microinteractions in Mobile and Wearable Computing. In *Advanced Information Networking and Applications (AINA), 2016 IEEE 30th International Conference on* (pp. 495-500). IEEE.

Baasanjav, U. (2013). Incorporating the experiential learning cycle into online classes. *Journal of Online Learning and Teaching, 9*(4), 575.

Bajaj, J., Harlalka, A., Kumar, A., Punekar, R. M., Sorathia, K., Deshmukh, O., & Yadav, K. (2015). Audio Cues: Can Sound Be Worth a Hundred Words?. In *Learning and Collaboration Technologies* (pp. 14-23). Springer, Cham.

Berman, R. A., & Colby, C. L. (2002). Auditory and visual attention modulate motion processing in area MT+. *Cognitive Brain Research*, *14*(1), 64-74.

Beard, C., Wilson, J. P., & McCarter, R. (2007). Towards a Theory of E-learning: Experiential e-learning. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *6*(2), 3-15.

Bindra, D. (1978). How adaptive behavior is produced: a perceptual-motivational alternative to response reinforcements. *Behavioral and Brain Sciences*, *1*(1), 41-52.

Bishop, M. J. (2000). The systematic use of sound in multimedia instruction to enhance learning (pp. 1-156). Lehigh University.

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.

Bishop, M. J., & Sonnenschein, D. (2012). Designing with sound to enhance learning: four recommendations from the film industry. *Journal of Applied Instructional Design, 2*(1), 5-15.

Boltz, M., Schulkind, M., & Kantra, S. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition, 19*(6), 593-606.

Borchers, J. O. (2000, August). A pattern approach to interaction design. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 369-378). ACM.

Brewster, S. A. (1997). Using non-speech sound to overcome information overload. *Displays, 17*(3-4), 179-189.

Brewster, S. A. (2002). Non-speech auditory output. *The human-computer interaction handbook,* 220-239.

Brophy, J. (2010). *Motivating students to learn* (3rd ed.). New York, NY: Routledge.

Brown, M. L., Newsome, S. L., & Glinert, E. P. (1989, March). An experiment into the use of auditory cues to reduce visual workload. In *ACM SIGCHI Bulletin* (Vol. 20, No. SI, pp. 339-346). ACM.

Buxton, W., Gaver, W., & Bly, S. (1994). Auditory interfaces: The use of non-speech audio at the interface. *Unpublished manuscript.*

Buxton, B. (2010). *Sketching user experiences: getting the design right and the right design.* Morgan Kaufmann.

Buxton, W. (1989). Introduction to this special issue on non-speech audio. *Human-Computer Interaction, 4,* 1-9.

Calandra, B., Barron, A. E., & Thompson-Sellers, I. (2008). Audio use in e-learning: What, why, when, and how?. *International Journal on ELearning*, *7*(4), 589.

Calvert, S. L. (1981). The Effects of Selective Attention to Television Forms on Children's Comprehension of Content.

Calvert, S. L., & Gersh, T. L. (1985). Developmental differences in children's television story comprehension: Effects of content cues and auditory formal production features.

Calvert, S. L. (1985). Sound Effects and Content Cues for Children's Television Story Comprehension.

Calvert, S. L., & Scott, M. C. (1987). Sound Effects for Children's Comprehension of Variably-Paced Television Programs.

Calvert, S. L., & Scott, M. C. (1989). Sound effects for children's temporal integration of fast-paced television content. *Journal of Broadcasting & Electronic Media*, *33*(3), 233-246.

Carver, R. (1996). Theory for practice: A framework for thinking about experiential education. *Journal of Experiential Education*, *19*(1), 8-13.

Carver, R., King, R., Hannum, W., & Fowler, B. (2007). Toward a model of experiential e-learning. *MERLOT Journal of Online Learning and Teaching*, *3*(3), 247-256.

Cereijo, M. V. P. (2006). Attitude as predictor of success in online training. *International Journal on E-Learning*, *5*(4), 623-639.

Chen, K.-C., & Jang, S.-J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior*, *26*(4), 741-752.

Chou, P. T. M. (2010). Attention Drainage Effect: How Background Music Effects Concentration in Taiwanese College Students. *Journal of the Scholarship of Teaching and Learning*, *10*(1), 36-46.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Csapó, Á., & Wersényi, G. (2013). Overview of auditory representations in human-machine interfaces. *ACM Computing Surveys (CSUR)*, 46(2), 19.

Danna, J., Fontaine, M., Paz-Villagrán, V., Gondre, C., Thoret, E., Aramaki, M & Velay, J. L. (2015). The effect of real-time auditory feedback on learning new characters. *Human Movement Science, 43*, 216-228.

Daniels, L. (1995). Audio-Vision: Audio-Visual Interaction in Desktop Multimedia.

Daniels, H. L., & Moore, D. M. (2000). Interaction of cognitive style and learner control in a hypermedia environment. *International Journal of Instructional Media, 27*(4), 369.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology, 18*(1), 105.

Domagk, S., Schwartz, R. N., & Plass, J. L. (2010). Interactivity in multimedia learning: An integrated model. *Computers in Human Behavior*, 26(5), 1024-1033.

Dyer, J. F., Stapleton, P., & Rodger, M. (2017). Mapping Sonification for Perception and Action in Motor Skill Learning. *Frontiers in neuroscience, 11*, 463.

Epstein, A. N. (1982). Instinct and motivation as explanations for complex behavior. In *The physiological mechanisms of motivation* (pp. 25-58). Springer, New York, NY.

Fentress, J. C. Dynamic boundaries of patterned behaviour: Interaction and self-organization. In: Bateson, P.P.G.; Hinde, R.A., eds. Growing Points in Ethology. Cambridge: Cambridge University Press; 1976:135–169.

Ferati, M., Pfaff, M. S., Mannheimer, S., & Bolchini, D. (2012). Audemes at work: Investigating features of non-speech sounds to maximize content recognition. *International Journal of Human-Computer Studies, 70*(12), 936-966.

Franinović, K., & Salter, C. (2013). The experience of sonic interaction. *Sonic interaction design: Fresh perspectives,* 39-76.

Fumham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied cognitive psychology, 11*(5), 445-455.

Grice, S., & Hughes, J. (2009). Can music and animation improve the flow and attainment in online learning?. *Journal of Educational Multimedia and Hypermedia, 18*(4), 385.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and emotion, 24*(3), 175-213.

Guttentag, R. E. (1985). A developmental study of attention to auditory and visual signals. *Journal of Experimental Child Psychology, 39*(3), 546-561.

Hartnett, M. (2010). Motivation to learn in online environments: An exploration of two tertiary education contexts (Doctoral thesis). Massey University, Palmerston North, New Zealand. Retrieved from http://muir.massey.ac.nz/handle/10179/2043

Hartnett, M., George, A. S., & Dron, J. (2011). Examining motivation in online distance learning environments: Complex, multifaceted and situation-dependent. The *International Review of Research in Open and Distributed Learning, 12*(6), 20-38.

Hassenzahl, M. (2007). The hedonic/pragmatic model of user experience. *Towards a UX manifesto, 10*.

Hativa, N., & Reingold, A. (1987). Effects of audiovisual stimuli on learning through microcomputer-based class presentation. *Instructional Science, 16*(3), 287-306.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

Hermann, T., & Hunt, A. (2011). *The sonification handbook* (pp. 399-425). J. G. Neuhoff (Ed.). Berlin, Germany: Logos Verlag.

Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of experimental psychology: Applied, 11*(3), 157.

Hogan, M. J., Kelly, C. A., & Craik, F. I. (2006). The effects of attention switching on encoding and retrieval of words in younger and older adults. *Experimental Aging Research, 32*(2), 153-183.

Hupp, S. D., & Allen, K. D. (2005). Using an audio cueing procedure to increase rate of parental attention during parent training. *Child & family behavior therapy, 27*(2), 43-49.

Jørgensen, K. (2009). *A comprehensive study of sound in computer games: How audio affects player action*. Lewiston: Edwin Mellen Press.

Kerlinger, F. N. (1970). *A social attitude scale: Evidence on reliability and validity. Psychological Reports, 26*(2), 379-383.

Kerner, M. M. (1989). The art of the sound effects editor. Focal Press.

Kolb, D. (1984). *Experiential Learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice-Hall

Kolb, D. A. (2014). Experiential learning: Experience as the source of learning and development. FT press.

Königschulte, A. (2015). Sound as Affective Design Feature in Multimedia Learning--Benefits and Drawbacks from a Cognitive Load Theory Perspective. *International Association for Development of the Information Society*.

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N & Evreinov, G. (1999). The sonification report: Status of the field and research agenda. Report prepared for the national science foundation by members of the international community for auditory display. *International Community for Auditory Display (ICAD), Santa Fe, NM.*

Laugwitz, B., Held, T., & Schrepp, M. (2008, November). Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 63-76). Springer, Berlin, Heidelberg.

Leavy, P. (2017). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches.* Guilford Publications.

Liu, C., White, R. W., & Dumais, S. (2010, July). Understanding web browsing behaviors through Weibull analysis of dwell time. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 379-386). ACM.

Lunney, D., & Morrison, R. C. (1990, August). Auditory presentation of experimental data. In *Extracting Meaning from Complex Data: Processing, Display, Interaction* (Vol. 1259, pp. 140-147). International Society for Optics and Photonics.

Mainemelis, C., Boyatzis, R. E., & Kolb, D. A. (2002). Learning styles and adaptive flexibility: Testing experiential learning theory. *Management learning, 33*(1), 5-33.

Mann, R. E. (1979). The Effect of Music and Sound Effects on the Listening Comprehension of Fourth Grade Students.

Martens, R. L., Gulikers, J., & Bastiaens, T. (2004). The impact of intrinsic motivation on e-learning in authentic computer tasks. *Journal of Computer Assisted Learning, 20*(5), 368-376.

Markus, M. L. (1987). Toward a "critical mass" theory of interactive media: Universal access, interdependence and diffusion. *Communication research, 14*(5), 491-511.

Mayer, R., 2005. Cognitive theory of multimedia learning. In R.E.Mayer (Ed.), the Cambridge Handbook of Multimedia Learning. Cambridge University Press, Cambridge, pp. 31-48.

Mayer, R. E., & Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 19.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of educational psychology, 93*(1), 187.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist, 38*(1), 43-52.

McMillan, S. J. (2002). Exploring models of interactivity from multiple research traditions: Users, documents, and systems. *Handbook of new media, 2*, 205-229.

Mihalca, L., & Miclea, M. (2007). Current trends in educational technology research. *Cognitie, Creier, Comportament/Cognition, Brain, Behavior*, *11*(1).

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive brain research, 14*(1), 115-128.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational psychology review*, *19*(3), 309-326.

Murphrey, T. P. (2010). A CASE STUDY OF EELEARNING: Using Technology to Create and Facilitate Experiential Learning. *Quarterly Review of Distance Education*, *11*(4).

O'brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychological bulletin*, *97*(2), 316.

Perham, N., & Vizard, J. (2011). Can preference for background music mediate the irrelevant sound effect?. *Applied Cognitive Psychology*, *25*(4), 625-631.

Perham, N., & Currie, H. (2014). Does listening to preferred music improve reading comprehension performance?. *Applied Cognitive Psychology*, *28*(2), 279-284.

Pfeiffer, D. G. (2008). Listen and learn: An investigation of sonification as an instructional variable to improve understanding of complex environments. *Computers in Human Behavior*, *24*(2), 475-485.

Riedel, E., Endicott, L., Wasescha, A., & Goldston, B. (2007). Continuous, interactive, and online: A framework for experiential learning with working adults. *Innovate: Journal of Online Education*, *3*(6).

Rocchesso, D., Serafin, S., & Rinott, M. (2013). Pedagogical approaches and methods. *Sonic interaction design*, 125-150.

Rovai, A. P., Ponton, M., Wighting, M., & Baker, J. (2007). A comparative analysis of student motivation in traditional classroom and e-learning courses. *International Journal on E-Learning*, *6*(3), 413-432.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional science*, *23*(5), 405-431.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Saffer, D. *Microinteractions: Designing with Details.* O’Reilly, Sebastopol, CA, 2013.

Schaffert, N., Mattes, K., Barrass, S., & Effenberg, A. O. (2009, December). Exploring function and aesthetics in sonifications for elite sports. In *Proceedings of the 2nd international conference on music communication science* (ICoMCS2) (Vol. 83, p. 86). HCSNet.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a Benchmark for the User Experience Questionnaire (UEQ). *IJIMAI, 4*(4), 40-44.

Shroff, R. H., Vogel, D., Coombes, J., & Lee, F. (2007). Student e-learning intrinsic motivation: A qualitative analysis. Communications of the Association for *Information Systems, 2007*(19), 241-260.

Shroff, R. H., Vogel, D. R., & Coombes, J. (2008). Assessing individual-level factors supporting student intrinsic motivation in online discussions: A qualitative study. *Journal of Information Systems Education, 19*(1).

Smus, B. (2013). *Web Audio API: Advanced Sound for Games and Interactive Apps.* " O'Reilly Media, Inc.".

Spence, C., Ranson, J., & Driver, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Perception & psychophysics, 62*(2), 410-424.

Toates, F. (1997). The interaction of cognitive and stimulus–response processes in the control of behaviour. *Neuroscience & Biobehavioral Reviews, 22*(1), 59-83.

Trevitte, C., & Eskow, S. (2007). Reschooling society and the promise of ee-learning: An interview with Steve Eskow. *Innovate: Journal of Online Education, 3*(6).

Vickers, P., & Alty, J. L. (2002). Using music to communicate computing information. *Interacting with computers, 14*(5), 435-456.

Walker, B. N., Nance, A., & Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus. Georgia Institute of Technology.

Walker, B. N., & Nees, M. A. (2011). Theory of sonification. *The sonification handbook*, 9-39.

Watson, P. (2016). Rules of thumb on magnitudes of effect sizes. Retrieved from http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize

Wersényi, G. (2010). Auditory representations of a graphical user interface for a better human-computer interaction. In *Auditory Display* (pp. 80-102). Springer, Berlin, Heidelberg.

Xie, K., DeBacker, T. K., & Ferguson, C. (2006). Extending the traditional classroom through online discussion: The role of student motivation. *Journal of Educational Computing Research*, *34*(1), 67-89.

APPENDIX A SYSTEMATIC REVIEW SEARCH TERMS

ERIC	TI ((sound effect*) or (audio* w5 (stimuli or cue*))) OR AB ((sound effect*) or (audio* w5 (stimuli (or cue*))) OR SU "Noise (Sound)" OR SU "sound effects" or SU "Auditory Stimuli") AND (TI ((attention or retention or comprehension or attainment)) OR AB ((attention or retention or comprehension or attainment)) OR SU "Attention" OR SU "comprehension")	489 Retrieved
PsycINFO	(DE "Education" OR DE "Adult Education" DE "Curriculum" OR DE "Distance Education" DE "Elementary Education" DE "High School Education" OR DE "Higher Education" OR DE "Middle School Education" DE "Personnel Training" OR DE "Private School Education" OR DE "Public School Education" OR DE "Secondary Education" DE "Teacher Education" OR DE "Teaching" OR DE "Instructional Media" OR DE "Training" DE "Computer Training" OR AB education or training or curriculum) AND (((DE "Attention" OR DE "Attentional Capture" OR DE "Divided Attention" OR	127 Retrieved

	DE "Focused Attention" OR DE "Sustained Attention") OR (DE "Retention")) OR (DE "Comprehension") OR (TI ((attention or retention or comprehension or attainment)) OR AB ((attention or retention or comprehension or attainment))) AND (DE "Auditory Stimulation" OR DE "Auditory Displays" OR DE "Auditory Feedback" OR TI (("sound effect" or "sound effects") or (audio* w5 (stimuli or cue*))) OR AB (("sound effect" or "sound effects") or (audio* w5 (stimuli (or cue*))))	
Academic Search Complete	(AB (education or training or curriculum)) AND ((TI ((attention or retention or comprehension or attainment)) OR AB ((attention or retention or comprehension or attainment))) AND (TI (("sound effect" or "sound effects") or (audio* w5 (stimuli or cue*))) OR AB (("sound effect" or "sound effects") or (audio* w5 (stimuli (or cue*))))	233 Retrieved

APPENDIX B EIGHT CHARACTERS IN THE PRACTICE SECTION

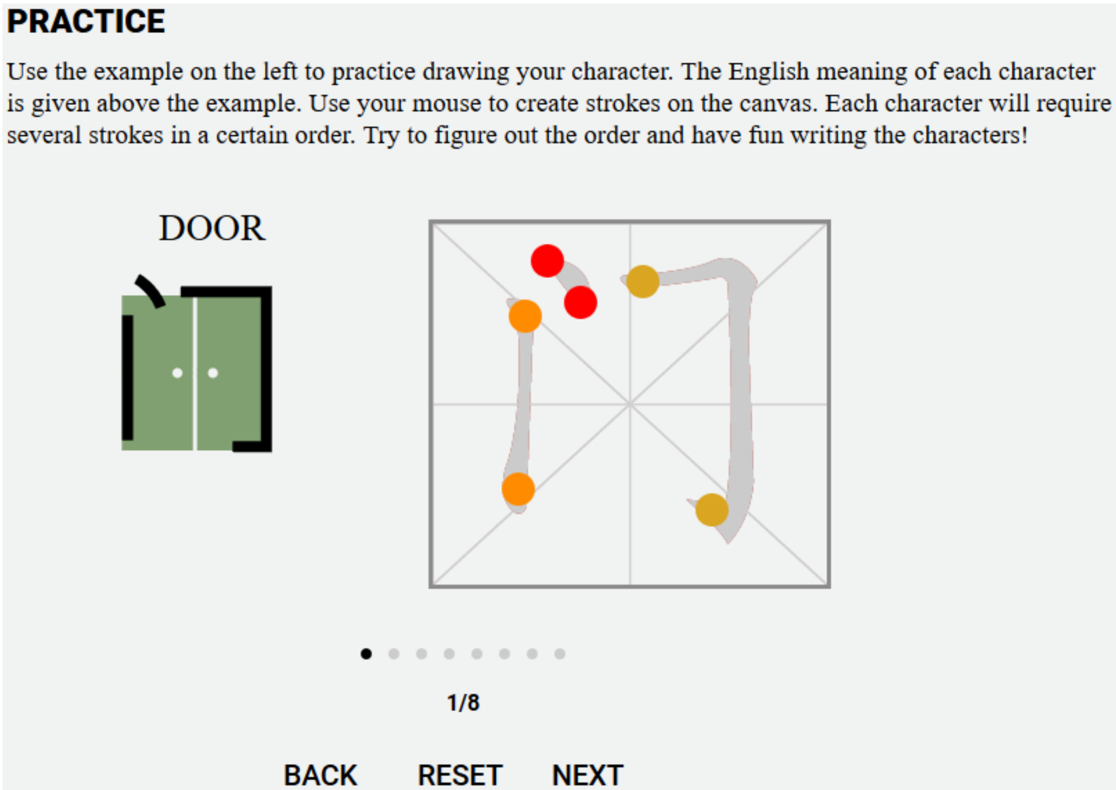


Figure 17. Character Door in the Practice Section

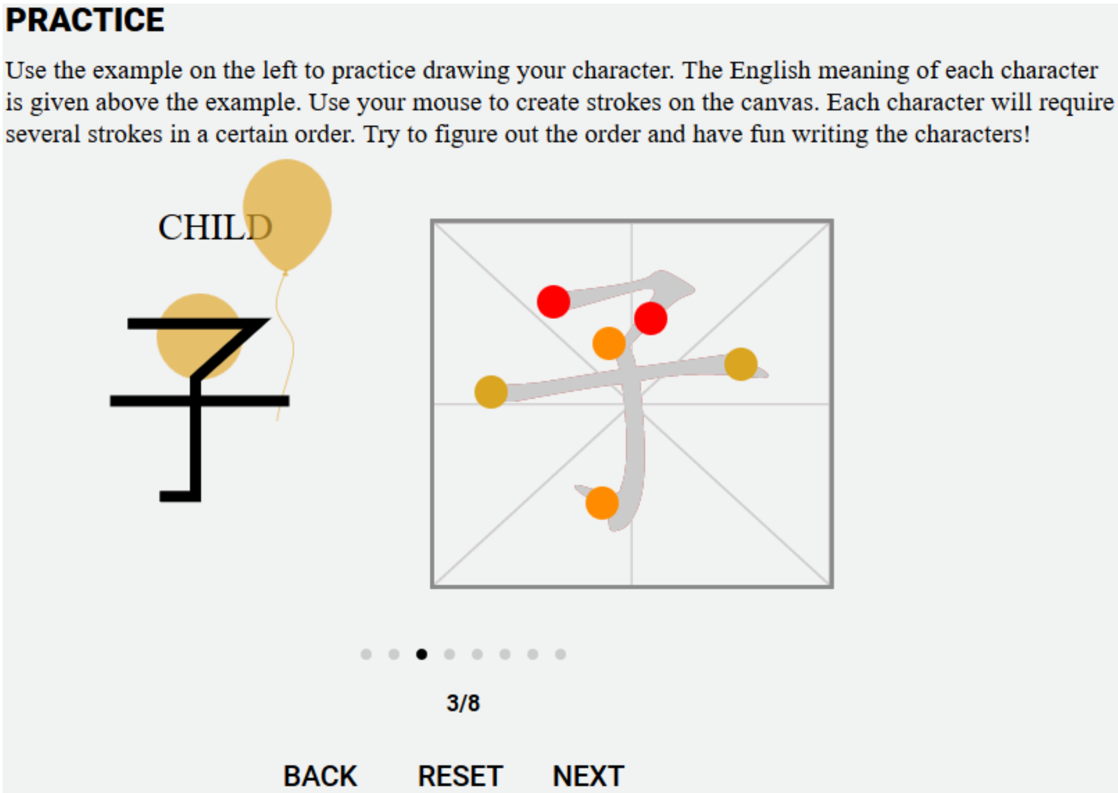


Figure 18. Character Child in the Practice Section

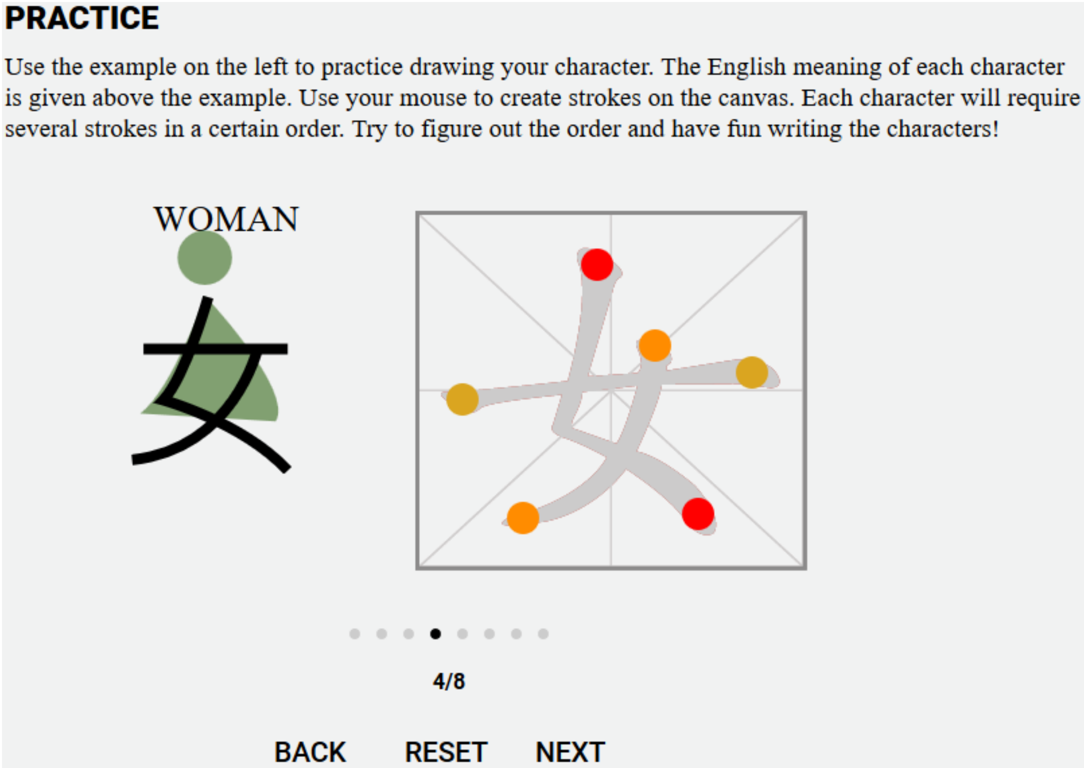


Figure 19. Character Woman in the Practice Section

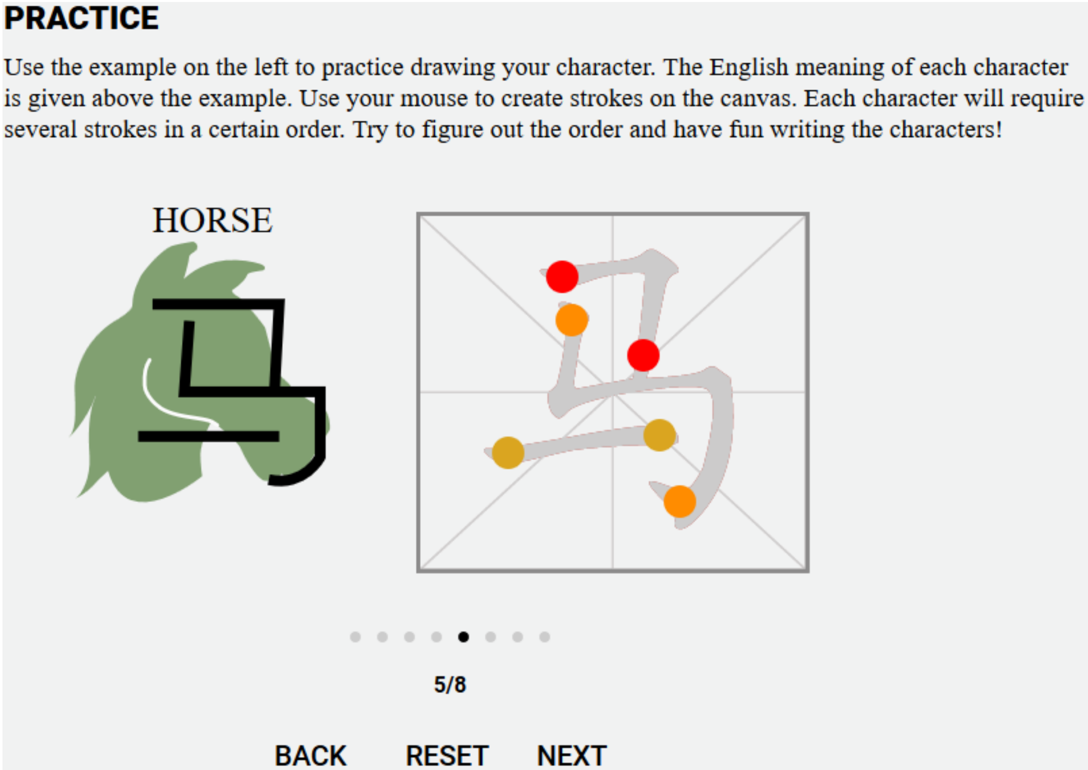


Figure 20. Character Horse in the Practice Section

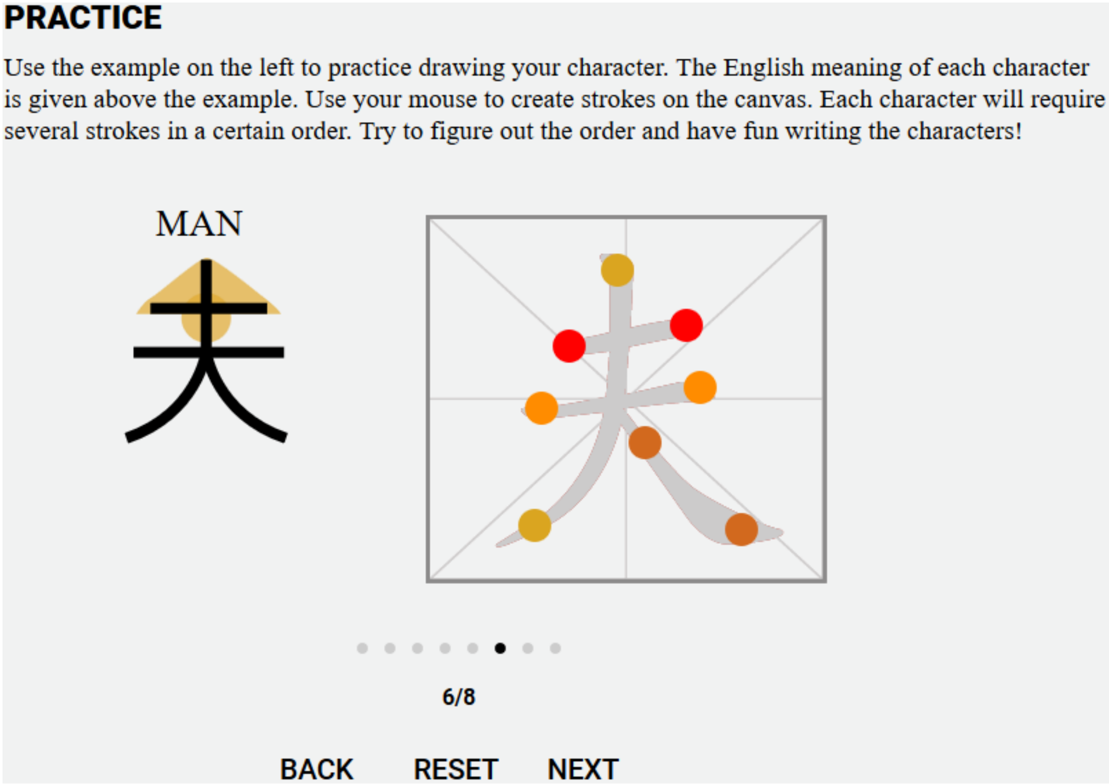


Figure 21. Character Man in the Practice Section

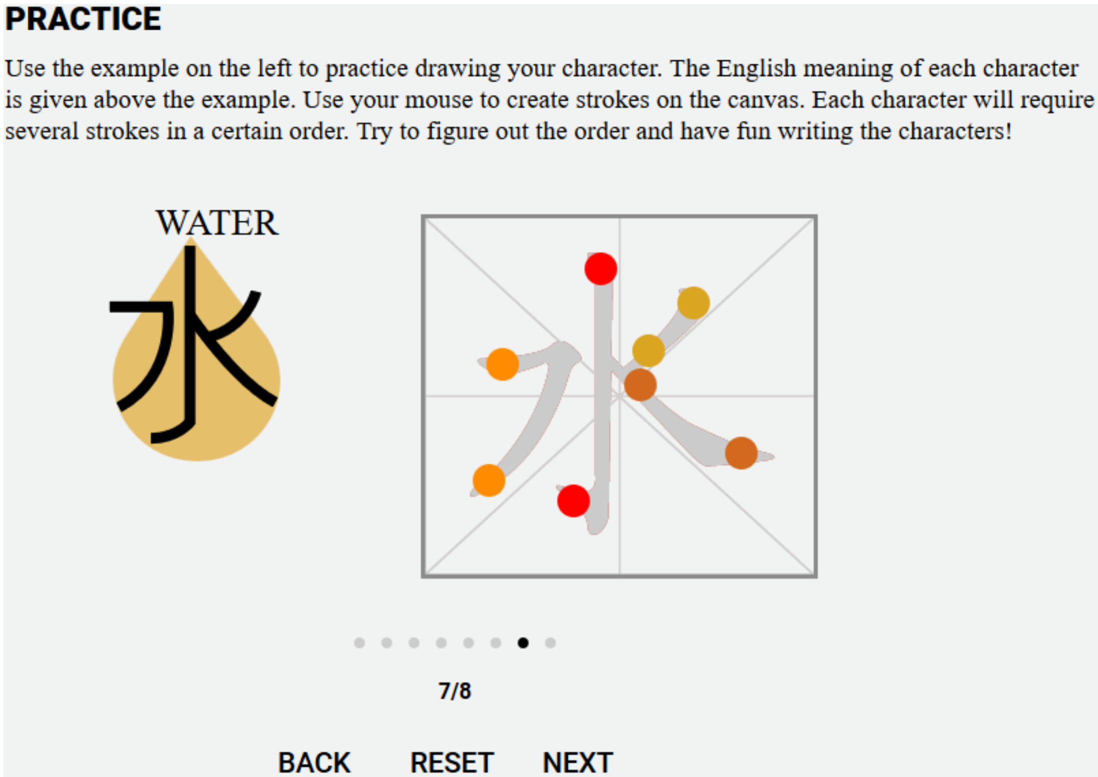


Figure 22. Character Water in the Practice Section

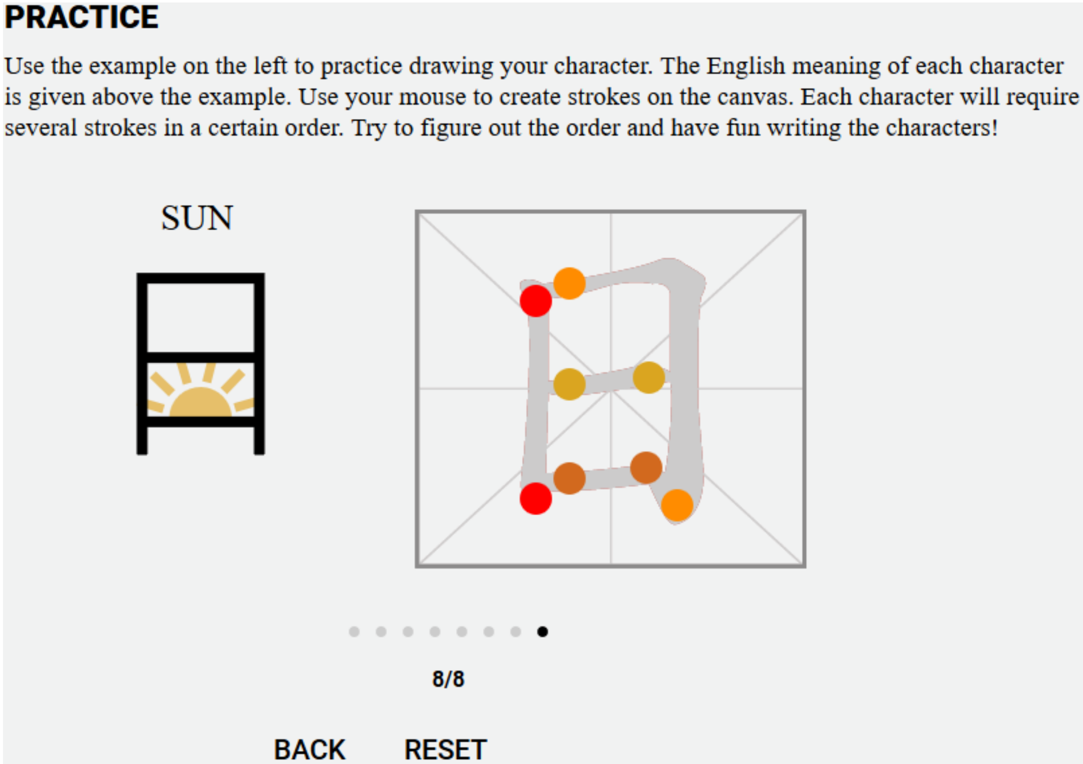


Figure 23. Character Sun in the Practice Section

APPENDIX C LEARNER CHARACTERISTICS QUESTIONS

We would like to know about you and your past online learning experience. Please answer each of the questions below as accurately as you can.

	What learning styles do you perceive yourself as?
○	Visual learner (i.e. prefer to learn through seeing drawings, pictures, and other image-rich teaching tools)
○	Auditory learner (i.e. prefer to learn by listening to lectures, exploring materials through discussions, and talking through ideas
○	Kinesthetic learner (i.e. prefer to learn through touching and experiences that emphasize doing, physical involvement, and manipulation of objects
	Have you had any online learning experience?
○	No
○	Yes, a little
○	Yes, a lot
	In general, what kind of experience was it?
○	Positive
○	Neutral
○	Negative

APPENDIX D USER EXPERIENCE QUESTIONNAIR (UEQ)

The following questionnaire consists of pairs of contrasting attributes that may apply to the Learning Chinese website. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by clicking the circles that most closely reflects your impression. Please decide spontaneously. Don't think too long about your decision but stay with your original impression. Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the Learning Chinese website. Nevertheless, please choose a circle in every line. It is your personal opinion that counts. Please remember: there is no wrong or right answer!

	1	2	3	4	5	6	7	
annoying	○	○	○	○	○	○	○	enjoyable
not understandable	○	○	○	○	○	○	○	understandable
creative	○	○	○	○	○	○	○	dull
easy to learn	○	○	○	○	○	○	○	difficult to learn
valuable	○	○	○	○	○	○	○	inferior
boring	○	○	○	○	○	○	○	exciting
not interesting	○	○	○	○	○	○	○	interesting
unpredictable	○	○	○	○	○	○	○	predictable

fast	○	○	○	○	○	○	○	slow
inventive	○	○	○	○	○	○	○	conventional
obstructive	○	○	○	○	○	○	○	supportive
good	○	○	○	○	○	○	○	bad
complicated	○	○	○	○	○	○	○	easy
unlikable	○	○	○	○	○	○	○	pleasing
usual	○	○	○	○	○	○	○	leading edge
unpleasant	○	○	○	○	○	○	○	pleasant
secure	○	○	○	○	○	○	○	not secure
motivating	○	○	○	○	○	○	○	demotivating
meets expectations	○	○	○	○	○	○	○	does not meet expectations
inefficient	○	○	○	○	○	○	○	efficient
clear	○	○	○	○	○	○	○	confusing
impractical	○	○	○	○	○	○	○	practical
organized	○	○	○	○	○	○	○	cluttered
attractive	○	○	○	○	○	○	○	unattractive
friendly	○	○	○	○	○	○	○	unfriendly

conservative ○ ○ ○ ○ ○ ○ ○ innovative

APPENDIX E THE SITUATIONAL MOTIVATION SCALE (SIMS)

Please click a circle that best describes the reason why you are engaged in Practice, Matching, and Stroke Order activities in the Learning Chinese website.

	Not at all	A very little	A little	Moderately	Enough	A lot	exactly
I think the activities are interesting	○	○	○	○	○	○	○
I am doing the activities for my own good	○	○	○	○	○	○	○
I think that these activities are pleasant	○	○	○	○	○	○	○
I think these activities are good for me	○	○	○	○	○	○	○
These activities are fun	○	○	○	○	○	○	○
I am engaged in these activities by personal decision	○	○	○	○	○	○	○
I feel good when doing these activities	○	○	○	○	○	○	○
I believe that these activities are important for me	○	○	○	○	○	○	○

APPENDIX F SPSS OUTPUT OF MANOVA

Box's Test of Equality of Covariance Matrices[a]

Box's M	102.588
F	1.026
df1	78
df2	2918.035
Sig.	.418

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + ID_Versions + ID_LS + ID_Attitude + ID_Versions * ID_LS + ID_Versions * ID_Attitude + ID_LS * ID_Attitude + ID_Versions * ID_LS * ID_Attitude

Figure 24. SPSS Output of Box's M

Bartlett's Test of Sphericity[a]

Likelihood Ratio	.000
Approx. Chi-Square	1014.643
df	5
Sig.	.000

Tests the null hypothesis that the residual covariance matrix is proportional to an identity matrix.

a. Design: Intercept + ID_Versions + ID_LS + ID_Attitude + ID_Versions * ID_LS + ID_Versions * ID_Attitude + ID_LS * ID_Attitude + ID_Versions * ID_LS * ID_Attitude

Figure 25. SPSS Output of Bartlett's Test of Sphericity3

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[d]
Intercept	Pillai's Trace	.934	566.884[b]	3.000	121.000	.000	.934	1700.651	1.000
	Wilks' Lambda	.066	566.884[b]	3.000	121.000	.000	.934	1700.651	1.000
	Hotelling's Trace	14.055	566.884[b]	3.000	121.000	.000	.934	1700.651	1.000
	Roy's Largest Root	14.055	566.884[b]	3.000	121.000	.000	.934	1700.651	1.000
ID_Versions	Pillai's Trace	.159	7.619[b]	3.000	121.000	.000	.159	22.858	.986
	Wilks' Lambda	.841	7.619[b]	3.000	121.000	.000	.159	22.858	.986
	Hotelling's Trace	.189	7.619[b]	3.000	121.000	.000	.159	22.858	.986
	Roy's Largest Root	.189	7.619[b]	3.000	121.000	.000	.159	22.858	.986
ID_LS	Pillai's Trace	.091	1.280	9.000	369.000	.246	.030	11.522	.627
	Wilks' Lambda	.910	1.298	9.000	294.633	.237	.031	9.444	.520
	Hotelling's Trace	.099	1.312	9.000	359.000	.229	.032	11.810	.640
	Roy's Largest Root	.092	3.790[c]	3.000	123.000	.012	.085	11.371	.804
ID_Attitude	Pillai's Trace	.060	1.257	6.000	244.000	.278	.030	7.540	.490
	Wilks' Lambda	.940	1.261[b]	6.000	242.000	.276	.030	7.568	.492
	Hotelling's Trace	.063	1.266	6.000	240.000	.274	.031	7.594	.493
	Roy's Largest Root	.059	2.409[c]	3.000	122.000	.070	.056	7.228	.590
ID_Versions * ID_LS	Pillai's Trace	.050	.701	9.000	369.000	.708	.017	6.310	.348
	Wilks' Lambda	.950	.699	9.000	294.633	.709	.017	5.096	.278
	Hotelling's Trace	.052	.698	9.000	359.000	.711	.017	6.279	.347
	Roy's Largest Root	.046	1.871[c]	3.000	123.000	.138	.044	5.612	.475
ID_Versions * ID_Attitude	Pillai's Trace	.021	.857[b]	3.000	121.000	.465	.021	2.572	.232
	Wilks' Lambda	.979	.857[b]	3.000	121.000	.465	.021	2.572	.232
	Hotelling's Trace	.021	.857[b]	3.000	121.000	.465	.021	2.572	.232
	Roy's Largest Root	.021	.857[b]	3.000	121.000	.465	.021	2.572	.232
ID_LS * ID_Attitude	Pillai's Trace	.031	.435	9.000	369.000	.916	.010	3.915	.216
	Wilks' Lambda	.969	.430	9.000	294.633	.919	.011	3.134	.175
	Hotelling's Trace	.032	.425	9.000	359.000	.921	.011	3.828	.211
	Roy's Largest Root	.017	.715[c]	3.000	123.000	.545	.017	2.145	.199
ID_Versions * ID_LS * ID_Attitude	Pillai's Trace	.016	.322	6.000	244.000	.925	.008	1.929	.140
	Wilks' Lambda	.984	.319[b]	6.000	242.000	.926	.008	1.917	.139
	Hotelling's Trace	.016	.317	6.000	240.000	.928	.008	1.904	.138
	Roy's Largest Root	.013	.542[c]	3.000	122.000	.655	.013	1.626	.159

a. Design: Intercept + ID_Versions + ID_LS + ID_Attitude + ID_Versions * ID_LS + ID_Versions * ID_Attitude + ID_LS * ID_Attitude + ID_Versions * ID_LS * ID_Attitude

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

d. Computed using alpha = .05

Figure 26. SPSS Output of Multivariate Test

Levene's Test of Equality of Error Variances[a]

		Levene Statistic	df1	df2	Sig.
Online_Experience	Based on Mean	.871	14	123	.591
	Based on Median	.501	14	123	.929
	Based on Median and with adjusted df	.501	14	85.653	.927
	Based on trimmed mean	.829	14	123	.636
Learner_Motivation	Based on Mean	1.203	14	123	.282
	Based on Median	.946	14	123	.512
	Based on Median and with adjusted df	.946	14	96.050	.514
	Based on trimmed mean	1.174	14	123	.304
Time	Based on Mean	1.912	14	123	.031
	Based on Median	1.261	14	123	.241
	Based on Median and with adjusted df	1.261	14	68.491	.254
	Based on trimmed mean	1.740	14	123	.056

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ID_Versions + ID_LS + ID_Attitude + ID_Versions * ID_LS + ID_Versions * ID_Attitude + ID_LS * ID_Attitude + ID_Versions * ID_LS * ID_Attitude

Figure 27. SPSS Output of Levene's Test

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[d]
Corrected Model	Online_Experience	11780.151[a]	16	736.259	1.479	.118	.161	23.665	.850
	Learner_Motivation	2124.483[b]	16	132.780	1.719	.051	.183	27.496	.910
	Time	876854.309[c]	16	54803.394	.851	.626	.100	13.620	.551
Intercept	Online_Experience	531364.259	1	531364.259	1067.466	.000	.897	1067.466	1.000
	Learner_Motivation	46373.436	1	46373.436	600.194	.000	.830	600.194	1.000
	Time	6960625.794	1	6960625.794	108.117	.000	.468	108.117	1.000
ID_Versions	Online_Experience	3979.293	1	3979.293	7.994	.005	.061	7.994	.801
	Learner_Motivation	1129.547	1	1129.547	14.619	.000	.106	14.619	.967
	Time	54860.326	1	54860.326	.852	.358	.007	.852	.150
ID_LS	Online_Experience	4896.996	3	1632.332	3.279	.023	.074	9.838	.739
	Learner_Motivation	91.827	3	30.609	.396	.756	.010	1.188	.127
	Time	35007.110	3	11669.037	.181	.909	.004	.544	.083
ID_Attitude	Online_Experience	910.645	2	455.323	.915	.403	.015	1.829	.205
	Learner_Motivation	3.825	2	1.912	.025	.976	.000	.049	.054
	Time	418441.857	2	209220.929	3.250	.042	.050	6.500	.610
ID_Versions * ID_LS	Online_Experience	1090.765	3	363.588	.730	.536	.018	2.191	.202
	Learner_Motivation	290.635	3	96.878	1.254	.293	.030	3.762	.329
	Time	29460.543	3	9820.181	.153	.928	.004	.458	.077
ID_Versions * ID_Attitude	Online_Experience	961.961	1	961.961	1.932	.167	.015	1.932	.281
	Learner_Motivation	6.657	1	6.657	.086	.770	.001	.086	.060
	Time	21926.268	1	21926.268	.341	.561	.003	.341	.089
ID_LS * ID_Attitude	Online_Experience	990.812	3	330.271	.663	.576	.016	1.990	.187
	Learner_Motivation	45.759	3	15.253	.197	.898	.005	.592	.086
	Time	78385.459	3	26128.486	.406	.749	.010	1.218	.129
ID_Versions * ID_LS * ID_Attitude	Online_Experience	139.166	2	69.583	.140	.870	.002	.280	.071
	Learner_Motivation	14.162	2	7.081	.092	.912	.001	.183	.064
	Time	91927.657	2	45963.828	.714	.492	.011	1.428	.168
Error	Online_Experience	61227.070	123	497.781					
	Learner_Motivation	9503.488	123	77.264					
	Time	7918784.684	123	64380.363					
Total	Online_Experience	3314563.000	140						
	Learner_Motivation	270660.000	140						
	Time	60428285.000	140						
Corrected Total	Online_Experience	73007.221	139						
	Learner_Motivation	11627.971	139						
	Time	8795638.993	139						

a. R Squared = .161 (Adjusted R Squared = .052)

b. R Squared = .183 (Adjusted R Squared = .076)

c. R Squared = .100 (Adjusted R Squared = -.017)

d. Computed using alpha = .05

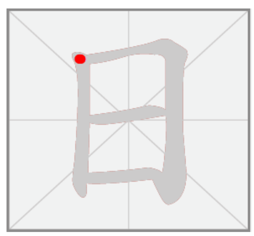
Figure 28. SPSS Output of Univariate ANOVA Tests

APPENDIX G INVITATION EMAIL

Invitation Email

Howdy! My name is Yun Li, a Ph.D. candidate specialized in Educational Technology. I would like to invite you to participate in a research study.

In this study, you will participate in an interactive learning experience to discover how the Chinese language has evolved. In addition, you will learn to write some characters. The graphic below shows an example of the writing canvas that you will be interacting with in this multimedia learning website.



(Sun)

Participation involves a 30 minutes online learning and approximately 5 minutes for a survey. You will participate in this study at your preferred location and time. If you are interested in having some interactive online learning experience or learning Chinese language, please click HERE to sign up this research study. For detailed information, please read the enclosed 'Information Sheet'.

Upon the completion of the Learning Chinese program, you will be placed in a lucky draw to win a gift card of $25, $50, or $100.

Yun Li

Ph.D. Candidate

Department of Educational Psychology

Texas A&M University

liyun215@tamu

APPENDIX H RESEARCH INSTRUCTION

Instruction for the Treatment Group

Howdy,

Thanks for being willing to participate in an online learning experience research study. In this study, you will be provided an interactive learning experience during which you will discover how the Chinese language has evolved. In addition, you will learn to write some Chinese characters on canvas. You will work on this website at you preferred location.

If you decide to participate, please read the instructions as below:

- **Please be prepared to use a computer/laptop.** The website doesn't function properly using mobile devices such as smartphones, iPad, and tablets.
- **Also turn on your audio but DO NOT use a headset.** Find a quiet location without distractions or environmental noise.
- **As a choice of internet browser,** please use either the Mozilla Firefox or Google Chrome, copy the link (www.dmliyun.com) and paste it in the browser address bar.
- **On the front page of the Learning Chinese website (**www.dmliyun.com**), you will need to type your name to confirm your participation.**
- **You will be assigned bonus points and be placed in a lucky draw to win a gift card once you complete the survey on the website.** You will find the survey in the Stroke Order section of the website. Once you completed the Stroke Order section, a message box will pop up with the survey link.

Upon the completion of the survey, your name will be recorded for the bonus points accordingly.

In addition, I will contact the winners of the lucky draw around the end of May 2018 by email.

Please note that the website will be open from **Monday, June 11th to Sunday, June 17th at 11:55 pm**. The website will be closed after Sunday, June 17th at 11:55 pm.

For any questions or support with technical issues, please contact me via liyun215@tamu.edu or 979-618-0564.

Yun Li l Ph.D. Candidate

Department of Educational Psychology

Texas A&M University

liyun215@tamu

IRB Number: IRB2018-0149M

IRB Approval Date: 03/01/2018

IRB Expiration Date: 02/28/2023

Instruction for the Control Group

Howdy,

Thanks for being willing to participate in an online learning experience research study. In this study, you will be provided an interactive learning experience during which you will discover how the Chinese language has evolved. In addition, you will learn to write some Chinese characters on canvas. You will work on this website at you preferred location.

If you decide to participate, please read the instructions as below:

- **Please be prepared to use a computer/laptop.** The website doesn't function properly using mobile devices such as smartphones, iPad, and tablets.
- **As choice of internet browser,** please use either the Mozilla Firefox or Google Chrome, copy the link (www.dmyunli.com) and paste it in the browser address bar.
- **On the front page of the Learning Chinese website (www.dmyunli.com), you will need to type your name to confirm your participation.**
- **Your name will be placed in a lucky draw to win a gift card once you complete the survey on the website.** You will find the survey in the Stroke Order section of the website. Once you completed the Stroke Order section, a message box will pop up with the survey link.

Upon the completion of the survey, your name will be recorded for the bonus points accordingly.

In addition, I will contact the winners of the lucky draw around the end of May 2018 by email.

Please note that the website will be open from **Monday, June 11th to Sunday, June 17th at 11:55 pm**. The website will be closed after Sunday, June 17th at 11:55 pm.

For any questions or support with technical issues, please contact me via liyun215@tamu.edu or 979-618-0564.

Yun Li l Ph.D. Candidate

Department of Educational Psychology

Texas A&M University

liyun215@tamu

IRB Number: IRB2018-0149M

IRB Approval Date: 03/01/2018

IRB Expiration Date: 02/28/2023