EFFECTS OF COGNITIVE STRESS IN HANDWRITING MOVEMENTS IN A PURSUIT LOOP-DRAWING TASK

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Jane Redfield Yank

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Thomas A. Stoffregen, Co-Adviser

Michael G. Wade, Adviser

November, 2010

Acknowledgements

The lyf so short, the craft so long to lerne….

Chaucer (1342/3-1400)

My journey to a doctoral degree had a long and unusual trajectory, marked by encounters with a fascinating array of talented scholars and researchers who generously aided, instructed, and guided me along the way, expanded my horizons, and led me far beyond my expectations.

First, I would like to acknowledge Herbert L. Pick, Jr. PhD, who was one of the first American researchers to investigate the role of perception and action in handwriting. When I balked at the challenge of pursuing a doctoral degree, Herb encouraged me, and later served as adviser during part of my program.

Michael Wade, PhD stepped in to serve as my adviser upon the untimely and tragic death of my first adviser, the beloved Allen W. Burton, PhD.

Thomas Stoffregen, PhD generously continued my advising when my progress stalled. Without his assistance, encouragement, and persistent nudges, I would not have completed this degree.

Arend Van Gemmert, PhD, Assistant Professor of Kinesiology at Louisiana State University, guided me through the lengthy process of developing and implementing my research project. Since few handwriting scientists conduct research in the United States, Dr. Van Gemmert's willingness to assist me through all the phases of the project was particularly invaluable and generous.

Gary Oehlert, PhD, Professor of Statistics and Associate Dean of the College of Liberal Arts at the University of Minnesota, provided invaluable guidance and careful critiques far beyond any reasonable expectations. His wit, insight, and patience added immeasurably to my morale and to the quality of this project.

Carol Leitschuh, PhD revealed to me the role of a scholar in the broader academic and cultural world beyond my department and my discipline, while providing greatly appreciated emotional support throughout the final phases of my program.

The wise and creative administrative staff of the School of Kinesiology at the University of Minnesota cheerfully bailed me out of one mishap after another as a student and instructor while showing me how to work successfully within the university system.

The faculty in the School of Statistics, including Dr. Oehlert, deserve special commendation for their uncommon willingness to work with this underprepared graduate student; their teaching skill and kindness motivated me to appreciate a wide range of approaches for understanding data, and provided some of the best learning experiences I had at the university. Among these, I would like to acknowledge Michael Hardy, PhD who tutored me throughout these courses, dealt with my panic, and calmly helped me explore a variety of statistical methods for the data in this study.

My research benefited greatly from the assistance of many others who shared their impressive skills and knowledge with me. They include

- The members of the International Graphonomics Society who generously included me in their conferences and offered suggestions over the past decade

- Stan Wagon, PhD, professor of Mathematics and Computer Science, Macalester College, St. Paul, Minnesota, who assisted me with programming the experimental stimulus in *Mathematica*®, and invited me to participate in a week-long training program in *Mathematica*® programming

- Elizabeth Newton, PhD who assisted me with statistical programming in R

- James Haile, PhD, who instructed me in the methods of technical writing

- Greg Zarow, PhD, who taught me how to construct and defend scientific arguments.

In particular, I want to acknowledge the generous assistance of David Bailey, PhD of David Bailey Consultancy in Cheshire, UK, whose questions and suggestions taught me how science is conducted at the bench. He challenged me to think about how research questions could be translated into mathematical relationships, and how to use that knowledge to build programs for data analyses in *Mathematica*®. I deeply appreciate the insightful conversations we shared.

Finally, I want to express my appreciation to Alain Zuur, PhD who guided my explorations into methods of mixed model analysis.

Each of these individuals is a model teacher, and I am very grateful for the skills and information they so generously shared with me.

Dedication

In memory of my parents for displaying pride in my work and encouraging my studies

and

to my family - Larry, Joshua, Matthew, and Krista - for all the sacrifices you made, your support, and your humorous way of reminding me to keep things in perspective.

Your love made all the difference.

Abstract

The effects of concurrently presented visual traces of hand movements on timing, smoothness, and spatial accuracy were investigated during a loop-drawing tracking task. Thirty-five healthy young right-handed adults used a stylus on a digitizing tablet to track a left-to-right loop-drawing animation presented on a computer monitor. A dot target moved over a template of twelve connected cursive letter *e's*, leaving a track as it drew over each loop in the series. Participants were instructed to draw along with the target to reproduce the shape of the loops at the tempo of the target. Participants performed sixteen trials in a 2 × 2 design, eight trials with their trace visible on the computer monitor and eight trials without a visible trace, half with a constant target rate and half with a variable rate change mid-trial.

Spatial accuracy was greater when the participant trace was visible, as expected ($p < .0001$). An inverse relationship was found between drawing speed and spatial accuracy, consistent with the expectation that more spatial errors would occur at increased speeds. However, timing accuracy ($p < .0001$) and smoothness ($p = .0026$) decreased when the participant trace was visible. These results suggest that the visual trace of the participant tracking presented on the computer screen disrupted timing characteristics of perception-action coupling and increased the complexity of the task. Findings are discussed in the context of cognitive load.

TABLE OF CONTENTS

List of Tables

List of Figures

Number *Page*

CHAPTER I: INTRODUCTION

The origins of handwriting and drawing are rooted deep in antiquity. From early tokens and etchings, the skills of writing and drawing expanded over time into the development and use of increasingly sophisticated tools (Schmandt-Besserat, 1996). Handwriting and drawing skills have advanced civilization since its preliterate beginnings, extending knowledge, expressing views of nature and culture, providing records of human interaction, and enhancing human ability to think symbolically (Klein & Edgar, 2002). Despite the contributions that handwriting and drawing have made to human and technological advances over the millennia, however, many aspects of these abilities remain poorly understood. Those who wish to study them must take into account a variety of elements, including perception, haptic sense, coordination, kinematics, semantics, aesthetics, language, and memory. The connections among these elements imply that, in addition to their function as a method of recording and transmitting information, handwriting and drawing can provide a rich source of information about the operations of perceptually guided movement in the service of cognitive functions. Most recently, the study of handwriting has contributed to ongoing debates about the operations of perceptually guided movements, broadly defined, since findings from handwriting and drawing research may be useful for extending knowledge of coordinated behaviors to include cognitive elements (Stoffregen, personal communication, April, 2006).

In this introductory chapter, I will argue for the relevance of the study of handwriting, positioning it within the domain of cognitive psychology, particularly embodied cognition (Clark, 1999; Gibbs, 2005). Then I will outline the goals of this study, explain their significance, and present several hypotheses relevant to this work. Next, in Chapter 2, I will discuss the history of handwriting research and the theories that have guided its development, describing key propositions about spatial accuracy, timing, and stability from the competing theories of information processing and dynamical systems. Separately, I will describe research that has provided the foundation for the development of the target stimuli and the software for measuring the handwriting features relevant to this study. In Chapter 3, I will outline the experimental apparatus and methods used in this study. The statistical procedures will be presented in Chapter 4 with their rationale and applicability to the present data. In Chapter 5, I will outline the results of the experiment, and conclude with a discussion of these results and ideas for further study in Chapter 6.

1.1. Rationale for Studying Handwriting

Many educators, scholars, and writers of all ages have wondered why a study of handwriting is relevant in an age where written communications are dominated by keyboard technology (Thornton, 1996). Since handwriting instruction has been discontinued in many schools, and many adults bemoan the sorry state of their own handwriting (Pothier, 2007), a study of handwriting processes may seem hopelessly quaint. Dennis Baron (2009), Professor of English and Linguistics at the University of Illinois-Urbana is one who argues that handwriting is a nostalgic throwback to the past when handwritten text was more common than printed works. Florey (2009) disagrees, arguing for both keyboards and handwriting. She suggests that both writing tools reflect a "fascinating - and surprisingly crucial - 21st-century controversy" about differences in writing technologies and their influences on quality of thought and writing.

Declarations about the demise of handwriting are not new. Over the last century, many new communication devices, from telegraphs and telephones to ball point pens and word processors, have been heralded (or condemned, depending on one's point of view) as an invention that will eliminate handwriting (Thornton, 1996). Nevertheless, handwriting persists for several reasons. First, the tools for writing are many and accessible. Further, skilled handwriting has to be learned only once; after it is acquired, users apply the skill over a lifetime. In addition, handwriting adds a fuller dimension to personal messages, allowing greater immediacy between writers and readers than can be conveyed in a word-processed document (Haddock, 2007). Finally, in many situations, handwritten script and symbols are the most natural, efficient, and low-tech methods for recording and conveying information, even though keyboards have simplified some handwriting tasks and have become standard in many settings.

Compelling reasons for continuing to write by hand extend to the art of skillful communication as well. Scholars investigating writing, art, and gesture have found that handwriting may provide an advantage over keyboarding and texting for fostering improved communication, spelling, and language use (Richards, Berninger, & Fayol, 2009). For example, gesture has been observed to facilitate speech, as discovered in a study of gestures by speakers using the telephone (Bavelas, Gerwing, Sutton, & Prevost, 2008) and by those performing apparently meaningless gestures while speaking (Ravizza, 2003). Hence, if handwriting is a specialized form of gesture, as Gibson (1969) claimed, handwriting may facilitate speech and the cognitive structures used for communication. Recent evidence demonstrating the beneficial effects of handwriting on reading and compositional quality in elementary grades through high school supports this hypothesis (Berninger, Abbott, Augsburger, & Garcia, 2009; Connelly, 2009; Kandel, 2009; Shanahan, 2008). Studies showing the positive effect of writing on language have been supported by recent fMRI studies where researchers observed that more brain centers related to language were activated during handwriting than during keyboarding

(Connelly, 2009; Richards, Berninger, & Fayol, 2009). By linking handwriting to the shape of the human body and noting its effects on language, handwriting may be a specialized example of embodied cognition as well as a vehicle for recording and advancing human knowledge. Finally, handwriting enhances social elements of communication by allowing greater immediacy in personal messages between writers and readers than can be conveyed in a word-processed document or electronic text (Haddock, 2007).

Paradoxically, future methods of computing may increase the use of handwriting. Recognizing the naturalness and durability of speech, gestures, and handwriting for communication, groups of computer scientists are creating Natural User Interface (NUI) technologies, designed to use speech, gesture, and writing (i.e., pen-input devices) as computer inputs for many functions (Leopold & Ambler, 1997; NUI, 2009), making users much less dependent on keyboards. Computer scientists have already invented software for recognizing script written by the user in the air space in front of the monitor (Linderman, Lebedev, & Erlichman, 2009), which will significantly reduce the need for keyboarding. Developments such as NUIs will make handwriting skills increasingly relevant in the future, although there is little doubt that fussy handwriting styles of the past will be replaced by streamlined script that is simplified for ease of writing, reading, and computer recognition.

For these reasons, I contend that handwriting remains a viable skill, worthy of scholarly attention and the efforts of educators and students alike.

1.2. Handwriting: A Special Case of Embodied Cognition

1.2.1. The Imprint of the Body and Gesture on the Development of Writing

The history of handwriting is long and rich, evolving from primeval stone etchings, finger markings in clay, and tokens, to diverse and detailed handwriting systems performed throughout the world today. Although modern people associate handwriting most closely with language, the earliest relics of ancient writing reveal its association with the shape of the human body and human movement. Archaeologists and anthropologists have uncovered evidence that early counting and writing symbols were based on the shapes of fingers, hands, arms, and, in a few cases, the feet (Beltran, 1999; Frøkjær-Jenson, 1979; Montgomery, 2003; Sharpe & Van Gleder, 2006). Archaeologist Leroi-Gourhan (1993) proposed that many early marks found in prehistoric rock art symbols represented human gestures. He claimed that claimed that such marks codified gestures into permanent transmittable records of human activity and knowledge that fostered survival at a particular time and place. For example, Wildgen (2004) noted in *The Evolution of Human Language* that the success of the human species depended on successful adaptation to environmental affordances through knowledge that could be "systematically exploited, conserved, and transmitted" (p. 96) by symbols that would be understood by present and future generations. Thus, handwriting was not only a useful tool for communicating and recording current information, but essential for informing others about the affordances of the environment that would ensure their survival and growth.

1.2.2. Early Symbols and Psychological Development

The early symbols that evolved into handwriting may have provided an impetus to the development of human psychology as well. According to paleoanthropologist Richard Klein (2001), development of meaningful marks and symbols fostered a transition from prehistoric cognition to modern forms of thought and knowledge, characterized by self-awareness, group identity, and knowledge of the social and natural world. According to Klein's controversial view, symbolic mark-making is an essential mechanism for expanding cognitive capacity, extending knowledge, and creating social structures. Despite vigorous debates over the assertions of Leroi-Gourhan and Klein, scholars agree that the human form, gesture, language, cognition, and writing are mutually dependent (Arbib, 2000; Armstrong & Wilcox, 2007; Copple, 2003; Corballis, 2003; Kelly, Iverson, Terranova, Niego, Hopkins, & Goldsmith, 2002). Gibson's (1969) description of handwriting as "frozen gesture" has been cited in these debates, and supports the argument that handwriting is a special case of embodied cognition, a psychomotor system situated within the physical and social environment. As such, handwriting is recognized as an abundant source of information about the relationship between dynamic bodily characteristics, development, thought, and human expression within particular environments.

1.2.3. Renewed Interest in Handwriting Research

Over the last four decades, observations of handwriting have been used to contribute to ongoing debates about the nature of perceptually guided motion found in a universe of coordinated behaviors, leading to a resurgence of handwriting research. Handwriting studies have become a rich area of research, addressing cognitive psychology and motor development, perceptual-motor coordination, neuroscience, motor disorders, education and language, forensic sciences, and computer science. The results of these studies are strengthening the view that handwriting, as a specialized form of gesture, embodies and enhances cognitive processes, while providing unique insights into the physical, developmental, emotional, cognitive, and social characteristics of the writer.

1.2.4. Challenges for Handwriting Research

Handwriting is a challenging area of study. Unlike some other coordinated skills, handwriting involves a complex interplay of motor, environmental, and cognitive variables that the researcher must consider. These variables also interact with perception, haptic and tactile senses, proprioception, coordination, kinematics, and transient biomechanical and physiological features of the writer. In addition to dynamical features shared with other coordinated actions, handwriting also incorporates additional cognitive components, such as semantics, aesthetics, language, and memory. As an example of this complexity, the contents of written material have been shown to effect kinematics of writing as well as the coordination and movement strategies applied in the handwriting task (Harralson, Teulings, & Farley, 2008). Consequently, much of the current handwriting research is limited to studies of highly practiced behavior, such as signatures or simple forms such as loops and lines, performed under controlled conditions. Ultimately, this complex interplay of stable and transient factors involved in handwriting has contributed to the difficulty in establishing a uniform theory of handwriting (Van Galen, 1991), and makes the handwriting research a very complex area of study. Thus, in the present study, I investigate hand motions that are prominent features of handwriting forms, but lack semantic content, focusing on connected loops that were shaped as the cursive letter "e". This form is found commonly found in modern Western alphabets, Arabic, cursive Hebrew, cursive Cyrillic, and Sanskrit.

1.3. Goals of the Present Study

The purpose of this research was to test observations from dynamical systems and information processing theories about the role of concurrent visual feedback in tracking performance. In a set of experiments that varied both the feedback and the speed of tracking tasks, I explored components from each of these theories. Subjects in a drawing task manually tracked a moving target proceeding at either a constant or a changing speed, with and without concurrent visual feedback. My goal was to understand more about how concurrent visual feedback affects performance and to determine whether timing perturbations could destabilize the effects. In order to assess these influences, I measured participants' temporal coupling and spatial accuracy with the moving target and the smoothness of their writing lines.

My second goal in this study was to develop computer programs to quantify the spatial accuracy, timing relationships, and lack of smoothness (wiggle) in a line of writing. The present study introduces software produced in Mathematica ® to perform the following tasks:

(a) Display *y*-maxima and *y*-minima for matching spatial and temporal features of recorded samples with the target;

(b) Segment curves to compute the root mean square error (RMSE) between the target and the recorded samples; and

(c) Quantify alternating disturbances in the writing trace ("wiggles").

Since the software designed for this study has not been tested in other contexts, these aspects of the study were exploratory.

1.4. Significance of the Study

Research studies on coordinated movements, broadly defined, are of interest to handwriting researchers since they have implications for handwriting and drawing. Handwriting and drawing are activities that have been shown to share some of the same dynamical characteristics and origins that have been found in a large variety of other motor behaviors (e.g., Kay, Saltzman, Kelso, & Schöner, 1987). The results of this study may advance our knowledge of factors that affect the coordination and stability of the writing system, and they may help us to assess the relative contribution of vision to temporal accuracy, spatial accuracy, and smoothness in drawn trajectories. Such explorations may lead to greater understanding of the human sensorimotor system under normal conditions and conditions of stress or disease. Research on timing in hand tracking tasks and undulations in the writing line are especially interesting, since scientists exploring have observed variations in these components in normal writers (Mergl, Tigges, Schroter, Moller, & Hegerl, 1999) and in writers affected by stress, injury, illness, and chemical agents such as medications, drugs, or alcohol (e.g., Phillips, Ogeil, & Rogers, 2008; Watson, Jones, & Sharman, 1997). This line of research has been applied in a study of patients treated for schizophrenia, where disturbances in the recorded trace, or “wiggles” in continuous curves drawn by patients indicated subclinical adverse effects of psychotropic medications (Caligiuri, Teulings, Dean, Niculescu, & Lohr, 2009; Caligiuri, Teulings, Filoteo, Song, & Lohr, 2006). Further studies of spatial, timing, and “wiggle” characteristics would be useful in clinical settings to assess cognitive and sensory impairments in patients who are unable or reluctant to report these symptoms.

Forensic document examiners have long been interested in spatial accuracy and timing features in handwriting, in addition to smoothness (Widla, 1990) for identifying features consistent with copying, tracing, and forced writing. Recent studies have explored the timing and dynamics of handwriting motions under conditions of forgery (Harralson, Teulings, & Farley, 2007; Van Gemmert & Van Galen, 1996).

1.5. Research Questions

In the present study, I chose to focus on the effects of concurrent visual feedback during a set of pursuit loop-drawing tasks involving hand motions that are fundamental components of Western handwriting (Overvelde & Hulstijn, 2009). The following research questions were addressed in current study:

(1) Does concurrent visual feedback improve or hinder spatial accuracy between the participant loop-drawing response and the animated target?

(2) Will spatial accuracy and timing accuracy between the computer target and pen movements be negatively correlated in accordance with Fitts' Law?

(3) Does concurrent visual feedback improve or hinder timing accuracy between the participant loop-drawing response and the animated target?

(4) What is the effect of concurrent visual feedback on smoothness, measured by minor disturbances (wiggles) in the writing line?

(5) Does gender contribute to any observed differences in timing, spatial accuracy, or smoothness?

These questions will be addressed in detail in the following sections.

1.5.1. Effects of Concurrent Visual Feedback on Spatial Accuracy

Anecdotal evidence and research reports from differing theoretical perspectives have observed that visual information has a positive effect on spatial accuracy in tracking tasks (Guadagnoli & Kohl, 2001; Miall, 1996), and in copying tasks in general (Elliott, Garson, Goodman, & Chua, 1991). In other words, participants reproduce shapes more accurately when they can see their own tracings in relation to the shapes they are attempting to copy. Thus, I expect superior spatial matching in trials with visual feedback than without.

The expectations that spatial accuracy improves when individuals are able to observe their accuracy in copying tasks and that spatial accuracy degrades at increased speeds are not in dispute in the present study. Instead, I am using these expectations to assess the computer programs developed for the current research. If results that are consistent with previous findings are obtained, they may provide limited evidence of the reliability of the programs, pending rigorous testing of the subroutines for validity.

1.5.2. Effects of Concurrent Visual Feedback on Timing Accuracy

On the one hand, it is reasonable to believe that individuals are more likely to couple the timing of their movements to the frequency of an oscillating animated target when trackings and target are visible in real time (Hurley & Lee, 2006; Kawashima et al., 2000). Information processing theorists may explain that concurrent visual feedback enhances feedback and feedforward mechanisms for movement guidance and error correction (Adams, 1971; Adams & Goetz, 1973; Hinder, Tresilian, Riek, & Carson, 2008; Miall, 1996; Seidler, Noll & Thiers, 2004, Wulf & Prinz, 2001). Several dynamical systems theorists have found that vision fosters frequency coupling, but propose another explanation, suggesting instead that vision aids entrainment of the participant response to the rhythmicity of the external stimulus (Beek, Peper, & Daffertshofer, 2002; Bressler & Kelso, 2001; and Lopresti-Goodman, Richardson, Silva, & Schmidt, 2008.)

On the other hand, it is also reasonable to propose that concurrent visual feedback increases the difficulty of a tracking task by presenting two streams of dynamic visual information (i.e., from both target and the response). Increased task complexity and difficulty has been associated with slower movements and increased hesitation errors that may reduce timing accuracy (Marquardt, Gentz, & Mai, 1999; Swinnen, Schmidt, Nicholson, & Shapiro, 1990; Van Gemmert & Van Galen, 1996). Welford (1973) hypothesized that concurrent visual feedback increases task difficulty by creating a dual task, consisting of a motor task (drawing) and a cognitive task (concurrent visual self-monitoring), resulting in increased cognitive load. Van Gemmert (1997) found that increased cognitive load from dual tasks produced greater biomechanical stiffness and slower movement times, and thus, greater timing errors in tasks such as those in the present study. My goal is not to add to the debate on task complexity and cognitive load in tracking tasks, but to determine whether concurrent visual feedback improves or impedes timing accuracy in the tracking tasks performed in the present study.

1.5.3. Effects of Concurrent Visual Feedback on Smoothness ("Wiggles")

The effect of concurrent visual feedback was explored for its effects on smoothness of the recorded trace. Since few studies have addressed dysfluency in handwriting or drawing movements, a theoretical basis for predictions is limited. In one study of writing fluency, disturbances in the writing increased when the subjects' attention was directed to the visual trace of their pen motions (Marquardt, Gentz, & Mai, 1996). In another study investigating the relationship of handwriting disturbances to writing speed, Chen, Cha, Chee, and Tappert (2003) attributed increased "wrinkliness" in the writing line to slower writing speeds. Several research groups (Berryhill, Kveraga, Boucher, & Hughes, 2004; Kording & Wolpert, 2004) have related task complexity with greater uncertainty, producing slower movement speeds, and influencing the extent of “wiggles” observed in the recorded trace. Given the lack of research in this area, the results from this investigation on the relationship of concurrent visual feedback to the percentage of “wiggles” in the present study are exploratory.

1.5.4. Effect of movement speed on spatial accuracy

Increased movement speeds are expected to produce more spatial errors in accordance with Fitts' Law (Fitts, 1954), which predicts a speed-accuracy tradeoff in aiming tasks. The design of the present study conflates movement speed with other predictors (concurrent visual feedback, and in particular, changing target rates, so the exploration of this relationship is limited in this study to an experimental condition where the data are most reliable.

1.5.5. Effect of Gender on Handwriting

Current educators have observed that handwriting skill appears to be more advanced in females than in males, beginning at an early age and persisting over the lifespan (Hartley, 1991). There is disagreement, however, about whether these differences are innate or have socio-cultural roots (e.g., Burr 2002; Hayes, 1996). Der and Deary (2006) observed that males tend to write more quickly than females. As a result, one possibility is that males may demonstrate improved skill at matching their movements with the pace of the target in the present study than females. Other aspects of handwriting by contemporary male writers (decreased legibility and poorer form) may produce greater spatial errors and percentages of wiggle in the recorded trace.

In contrast, Mergl, Tigges, Schroter, Moller, and Hegerl (1999) found few gender differences in writing. Thus, another possibility is that few differences between genders will be observed in the present study. Thus, as a final exploratory question, I assessed differences between males and females on measures of spatial accuracy, timing accuracy, and smoothness in the recorded trace.

1.6. Research Hypotheses

1.6.1. Hypotheses for the Effect of Vision on Spatial Accuracy

Given the consensus on the benefits of vision for spatial accuracy in tracking (Elliot et al., 1991; Miall, 1996), visual feedback is expected to foster spatial accuracy; thus, spatial errors will be reduced in trials when the Participants' visual traces are presented concurrently on the monitor with the moving target.

(1) The effect of vision on spatial accuracy will be robust; a perturbation in the tempo of the target movement will not induce significant reduction in timing accuracy in samples where visual feedback is presented.

(2) Spatial accuracy will degrade at increased movement speeds in accordance with Fitts' Law (Fitts, 1954; Bosga, Meulenbroek & Rosenbaum, 2005; Zhai, Kong, & Ren, 2004).

Support for these hypotheses for spatial accuracy in the present study would be used to indicate that the software is reliable for measuring participant responses.

1.6.2. Hypotheses for the Effect of Vision on Timing Accuracy

(1) Visual feedback will foster timing accuracy; thus, timing errors will be reduced in trials when the Participants' visual traces are presented concurrently on the monitor with the moving target.

(1a) The effect of vision on timing accuracy will be robust; a perturbation in the tempo of the target movement will not induce significant reduction in timing accuracy in samples where visual feedback is presented.

(2) Alternately, the visual trace will have an adverse effect on timing accuracy, resulting in reduced timing accuracy when the Participants' visual traces are presented with the moving target. Timing accuracy will be improved when the visual trace is not presented on the screen.

(2a) In this case, a perturbation in the tempo of the target movement will further impede timing accuracy in samples where visual feedback is presented.

1.6.3. Hypotheses for the Effect of Vision on "Wiggles" in the Recorded Trace

(1) Visual feedback will foster smoother writing as demonstrated by a lower percentage of wiggles in samples drawn with visual feedback when compared to samples drawn without visual feedback.

(1a) The tempo of the target, whether moving at a constant or changing rate, will not affect this result.

(2) Alternately, the presence of the visual trace will disrupt the smoothness in the writing line, resulting in increased wiggles in samples drawn with visual feedback.

(2a) This effect will be robust to changing target rates.

1.6.4. Hypotheses for the Effect of Gender on Timing Accuracy

(1) Following Der and Deary (2006), gender differences in timing accuracy are expected, i.e., the timing patterns of males will result in fewer timing errors due to increased movement speeds.

(2) The effect of gender on timing will be robust; males will exhibit improved timing accuracy in trials where the recorded trace is presented (Feedback Present conditions) and where it is withheld (Feedback Absent conditions).

1.6.5. Hypotheses for the Effect of Gender on Spatial Accuracy and "Wiggles"

(1) Following Mergl et al. (1999), few gender differences are expected for measures of spatial accuracy or smoothness ("wiggles") in the writing line in any experimental condition (i.e., whether visual feedback is present or absent, and whether the target moves at a changing or constant rate across the sample).

(2) Alternately, writing facility has been observed to be superior in females than in males (Phelps & Stempl, 1987). Thus, greater spatial accuracy and smoothness may be found in females than in males. Since the number of males is small (n = 12) relative to females (n = 23) in the present study, results from this analysis will be interpreted with caution.

1.7. Methodological Limitations

Hypothesis testing requires the control of many variables, managed, in part, through randomization and structures within the research design. Nevertheless, an inevitable consequence of such control is a reduction in the naturalness or ecological validity of the experimental situation. This, in turn, raises questions about the extent to which findings from experimental research can generalize to behavior beyond the laboratory (Schmidt & Lee, 2995; Schmuckler, 2001). In the present study, handwriting differed from natural writing situations in several ways. Participants wrote on a digitizing tablet, using a non-inking stylus, and could view their recorded traces only on a video monitor that was displaced from the writing surface. Under these conditions of displaced feedback, the physical movements used in handwriting and the relation between vision and manual control will necessarily differ from ordinary handwriting. For these reasons, caution should be exercised in relating results of the present study to handwriting.

CHAPTER II: LITERATURE REVIEW

The present study is based on the premise that handwriting is a relevant area of research, capable of contributing to our understanding of human cognitive psychology and coordinated movement skills. Beginning with early psychological studies, this chapter describes the literature that supports this view. A discussion follows, focusing on the application of the most prominent theories of handwriting relevant to the current study, i.e., information processing and dynamical systems theories. The chapter concludes with a description of the research that was used to create the target stimuli, the selected independent and dependent variables, the computer programs developed to quantify the observed handwriting features, and the choice of methods used to analyze the study data.

2.1 History of Handwriting Research

2.1.1. Early Psychological Studies of Handwriting

Wilhelm Preyer (1841-1897), German neurophysiologist and founder of modern developmental psychology, pioneered the modern scientific study of the neurobehavioral factors in handwriting. In his 1895 text, *Zur Psychologie des Schreibens*, Preyer speculates about underlying neural control processes that might explain the fluid writing produced by individuals unable to write with their hands, but well-practiced in writing by mouth or feet. Later researchers, (e.g., Castiello & Stelmach, 1993; Merton, 1972; Meulenbroek, Rosenbaum, Thomassen, Loukopoulos, & Vaughan, 1996; and Raibert, 1977) updated and extended these studies, concurring with Preyer's early findings that well-practiced writing produced by other effectors, such as the non-dominant hand, foot, or the teeth, resembled the writing style of the dominant hand. More than a century after Preyer (1895), Wing (2000) replicated Preyer's observations through fMRI studies, discovering that similar patterns of brain activation occurred during signature writing, regardless of the effector chosen. Furthermore, he learned that fMRI patterns involved in writing were distinct from activation patterns seen when the effectors produced general non-writing motions, demonstrating that handwriting incorporated unique features not present in other motor skills.

The works of experimental psychologist June Downey (1875-1932), professor of psychology at the University of Wyoming, include many examples of research on the psychology of handwriting, with a focus on motor control processes. Between 1910 and 1931, she published more than twenty research articles and four scientific books on handwriting and mirror writing. As a researcher interested in the processes of motor control and learning, she focused upon the "interplay of visual, auditory, and kinaesthetic processes in the act of hand-writing" (Downey, 1910, p. 165-166). To achieve her objectives, she conducted studies of writing with the dominant and non-dominant hand over a range of conditions, varying the writing surfaces and introducing auditory and visual disturbances to the writing task. Her works contain meticulously detailed descriptions of the handwriting produced under these conditions.

After these early studies, interest in the psychological and motor aspects of handwriting waned. During the mid-1930s through the mid-1960s, medical and forensic journals continued to publish reports on handwriting changes accompanying diseases and the effects of chemicals on the central nervous system, but few psychological studies on skilled handwriting were conducted. Publications during this period included descriptions of handwriting by patients with Parkinson's disease (McLennan, Nakano, Tyler, and Schwab, 1972); multiple sclerosis, tremor, and mercury poisoning (Hirt, 1899); or by persons under the influence of alcohol (Hilton, 1969), hallucinogens, such as LSD (Hirsch, Jarvik, & Abramson, 1956), and drugs affecting the central nervous system (Dhawan, Bapat, & Saxena, 1969; Gross, 1975). Such handwriting studies were highly descriptive, but offered few examples of quantified measures, statistical analysis, or experimental research (Asicioglu & Turan, 2003). One notable exception was the work of Lewinson and Zubin (1942) which established a set of objective metric scales for evaluating the contraction and release of hand muscles occurring during rhythmic hand motions used in handwriting.

The emergence of cybernetics, defined as the scientific study of feedback and control mechanisms in animate and inanimate systems (Weiner 1948), rekindled interest in handwriting research. A new generation of scientists interested in control systems viewed handwriting as an accessible system for testing their theories. Denier van der Gon and Thuring (1965), for example, wrote one of the first widely distributed studies to use control systems technology for quantifying features of the handwriting process. In their study, participants were told to write connected looped *l*'s (i.e., *llll*) as quickly as possible on an iron table, using an iron pencil with a coil wrapped around the barrel. The coil allowed the researchers to increase the magnetic force and friction between the pen and the table. As the magnetic force increased, they noted that the size of the writing decreased. They estimated the time duration of samples written with different levels of magnetic force by counting dots comprising the written line that were generated by their pen and table writing system, an effort that was remarkable in its ingenuity and in the painstaking counting required. In addition to their study of the effects of magnetic forces, Denier van der Gon and Thuring (1965) analyzed the correlations between eye and hand accelerations during writing, using these findings to form hypotheses about the muscle innervations and the neurological functions invoked during the writing process.

By the end of the 1960's, handwriting attracted the interest of early cognitive psychology researchers who found handwriting tasks to be well-suited for investigating how information is processed in memory and language (Neisser, 1967). In a study involving figure-8 copying tasks, Goodnow & Levine (1973) linked cognitive and motor processes by observing that problem-solving strategies used by preschoolers, kindergarteners, first graders, and adults resembled their use of language rules, showing parallel trends in motor and language development.

Examples such as these handwriting studies from engineering and cognitive science showed that handwriting research not only was a promising vehicle for gaining insight into motor behavior, but also for exploring systems of language and cognition with associated mental representations and problem-solving strategies.

2.1.2. Contemporary Handwriting Research

Since these early beginnings, great strides have been made in handwriting and drawing research, particularly over the past three decades (Van Gemmert & Teulings, 2006). The science of handwriting and drawing, termed *graphonomics* (International Graphonomics Society, 1982), has prompted research from wide range of disciplines including medicine, psychiatry, neurology, psychology, education, forensic sciences, computing, engineering, and artificial intelligence. Scientists conducting graphonomics research have devised novel methods for studying complex motor skills, making significant contributions to overall knowledge of mechanisms for observing the processes of human movement. A selection of studies from several of these disciplines follows.

2.1.2.1. Recent Medical and Psychiatric Studies of Handwriting

Extensive research has been conducted on the writing of adults afflicted with cognitive and neurological disorders, with the greatest number of studies performed on handwriting in Parkinson's disease (Teulings, Contreras-Vidal, Stelmach, & Adler, 2002; Van Gemmert, Adler, & Stelmach, 2003). Distinct patterns of change have been identified in studies of handwriting in Huntington's disease (Phillips, Bradshaw, Chiu, & Bradshaw, 1994), multiple sclerosis (Longstaff & Heath, 2006), and Alzheimer's disease (Slavin, Phillips, Bradshaw, Hall, & Presnell, 1999; Werner, Rosenblum, Bar-On, Heinik & Korczyn, 2006). Other research has revealed characteristic features in the handwriting of adults with schizophrenia (Lange, Tucha, Aschenbrenner, Eichhammer, et al., 2001; Dose, Gruber, Grunz, Hook, Kempf, Scharfenberg, & Sick, 2007) and adult attention deficit disorder (Tucha, Tucha, Laufkötter, Mecklinger, Klein, & Lange, 2008). The handwriting of persons at risk of stroke has also shown distinct features (O'Reilly, Plamondon, Clément, Mathieu, & Lebrun, 2009).

In addition to the long-term changes in handwriting resulting from chronic or progressive diseases and aging (Dixon, Kurzen, & Friesen, 1993), temporary alterations in handwriting have been observed in adults experiencing the effects of untreated depression (Lange, Tucha, Aschenbrenner, Gottwald, et al., 2001; Lange, Tucha, Mecklinger, & Paul, 2003; Morrens, Wezenberg, Verkes, Hulstijn, Ruigt, & Sabbe, 2007; Mergl, et al., 2004). Other studies have observed the effects of chemical substances on handwriting, such as alcohol (Naito, Chang, & Shin, 2009; Tiplady, Baird, Lütcke, Drummond, & Wright, 2005), alcohol and cannabis (Zaki & Ibrahim, 1983), caffeine (Tucha, Stasik, Mecklinger, & Lange, 2006) nicotine (Tucha & Lange, 2004). Additional studies have observed handwriting changes with sleep deprivation (Jasper, Gordijn, Haüssler, Marquardt, & Hermsdörfer, 2009), changes in writing posture from standing, kneeling, and prone positions (Sciacca, Langlois-Peter, Gilhodes, Margot & Velay, 2009), and deception (Dick, Found, & Rogers, 2000; Luria & Rosenblum, 2009). Studies of children's handwriting have identified characteristic features in the writing of children with attention deficit disorder (Lange, Tucha, Walitza, et al., 2007), autism (Beversdorf, Anderson, Manning, Anderson, Nordgren, Felopulos, & Bauman, 2001), and developmental coordination disorder (Smits-Engelsman, Niemeijer, & van Galen, 2001).

While most of these studies were designed to provide only additional descriptive handwriting features for persons of different ages and conditions, some have been used to assess the effects of clinical interventions, pharmacological studies, and clinical trials of new medications. Recently, Caligiuri, Teulings, Dean, Niculescu, and Lohr (2007) and Caligiuri, Teulings, Filoteo, Song, and Lohr (2006) determined that irregular undulations or wiggles in the lines drawn by patients being treated for schizophrenia can be useful indicators of subclinical adverse effects of psychotropic medications.

2.1.2.2. Recent Psychological Studies of Handwriting.

Researchers investigating psycholinguistic, perceptual, and motor processes of handwriting and drawing owe a great debt to group of experimental and cognitive psychologists at Radboud University Nijmegen (Simner, Leedham, & Thomassen, 1996) who reinvigorated the field with experimental approaches and new technology. In 1986, they established the Center for Cognition (formerly, NICI, the Nijmegen Institute for Cognition and Information) to study action, intention, and motor control processes in spoken and written language. Since the founding of NICI, handwriting research has experienced a revival, fostering collaborations across disciplines to study handwriting factors in areas as diverse as cognition, motor behavior, motor disorders, education, language, artificial intelligence, forensic sciences, and software engineering (Meulenbroek and van Gemmert, 2003).

2.2 Development of Handwriting Theory

The central goal of all theories of motor control, including those pertaining to handwriting, is to explain how biomechanical, neurological, physiological, cognitive, and psychological components are constrained to produce smooth, effective movements. What sets handwriting apart from some other motor skills are psychological components, such as semantics, aesthetics, memory, cognition, language, stress, and emotion (Barrientos, 2002; Kandel, Álvarez, & Vallée, 2006; Lange, Tucha, Aschenbrenner, Gottwald, et al., 2001; Rosenblum, Chevion, & Weiss, 2006; Van Galen, 1991; Van Galen, Meulenbroek, & Hylkema, 1986; Van Gemmert, 1997; Zesiger, Mounoud, Hauert, 1993). As a result, not only is handwriting determined by physical laws that direct neuroanatomical systems of movement, it is influenced by cognitive processes within a broad psychological framework and the social and physical writing environment. Thus, an over-arching psychomotor theory of handwriting and drawing must incorporate multiple operations from a wide variety of influences.

At the same time, a comprehensive theory of handwriting must account for the stability and variability that characterize early and highly practiced handwriting and drawing motions (Newell & Van Emmerik, 1989). Stability in handwriting has been widely observed. Researchers concur that handwriting maintains a degree of structural integrity when writing conditions change, retaining enough similarities to be readily identified, in spite of multiple changing constraints on underlying cognitive and behavioral subsystems. Further, handwriting stability is robust, resisting forces that would completely degrade the process or the recognizability of the output, even when speed, size, and the visual trace from the writing are altered (Ihler, Fisher & Willsky, 2001; Wann & Nimmo-Smith, 1991a). On the other hand, researchers observe that handwriting is characterized by variability, agreeing with Teulings and Schomaker (1993) who noted that no characters in a sample of handwriting are likely to be identical, even when produced under ideal conditions.

An all-inclusive psychomotor theory of handwriting must therefore incorporate not only multiple perceptual, sensorimotor, psychological, and cognitive processes, but account for changes that occur with different task constraints, which alter the dynamics of handwriting motions with every change of utensil or surface. This task is made more challenging by the differing goals, perspectives, and methods of the various sciences conducting handwriting studies. While researchers in experimental psychology, neurobiology, perception, and motor control may seek answers to basic motor processes, others have different goals. For example, computer scientists and engineers may analyze variability to achieve more precise handwriting recognition, while cognitive scientists, linguists, and educators may be interested in learning about the cognitive architecture underlying language acquisition, expression, and literacy. Forensic scientists investigate the range of variability in writing, while medical scientists may explore changes in handwriting occurring with different physical conditions. Thus, the various approaches to graphonomics applied by diverse disciplines contribute to the difficulty in establishing any uniform theory of handwriting, even as they contribute to a broad perspective. Thus, van Galen (1991) remarked that the multiple levels of analysis, perspectives, methods, and goals in handwriting research preclude the possibility of a single inclusive handwriting theory, and that various handwriting theories will emerge to address specific physical, cognitive, linguistic, and social dimensions of handwriting.

2.3 Predominant Theoretical Views on the Psychology of Handwriting

At this time, information processing and dynamical systems theories are the predominant theoretical approaches to motor control processes in handwriting. Both theories have been applied to handwriting to account for motor variability (Glencross, 1980; Newell & Slifkin, 1998), complexity (Bernstein, 1967), stability (Getchell, 2006; Meulenbroek, Thomassen, van Lieshout, & Swinnen, 1998; Sallagoïty, Athènes, Zanone, & Albaret, 2004; Widla, 1990), and motor equivalence (Lashley, 1933; Meulenbroek, Thomassen, Rosebaum, Loukopoulos, & Vaughan, 1996; see Stelmach & Diggles, 1982, for a review).

Over one hundred years ago, Woodworth (1869-1962) (1899) addressed these issues by describing motor control as a stage process, in which central planning occurred prior to movement initiation and sensory feedback ensured movement accuracy. These ideas were later incorporated into information processing theories that developed in the 1950's (Miller, 1956). Concepts of motor programming grew out of this perspective to address the delays involved in moving through information processing stages (Schmidt, 1982), and have been widely applied to explain stability and automaticity in writing (Schmidt, 1975. For examples, see Castiello & Stelmach, 1993; Meulenbroek, Thomassen, van Lieshout, & Swinnen, 1998; Morasso, 1986; Stelmach, Mullins, & Teulings, 1984; and Teulings & Schomaker, 1993).

More recently, dynamical systems approaches have been applied to handwriting research. Briefly, dynamical systems theorists explain motor control as a function of softly assembled, multilevel coordinative structures emerging from properties of the task and individual, constrained by the environment, and subject to dynamical forces (Schaal, Mohajerian, and Ijspeert, 2007). Advances in computing and mathematical tools have made it possible for scientists to explore handwriting as a system with dynamical features such as self-organization and attractor states. These investigations have challenged traditional views, such as the role of pre-programming in handwriting.

Recent theoretical developments within information processing and dynamical systems perspectives hint at complementary roles for each approach that may advance the development of handwriting theory. For example, Van Soest and Van Galen (1995) theorized that applications of nonlinear dynamics to problems of handwriting may yield a fuller understanding of rapid state changes and noise within neural-musculo-skeletal models of handwriting. Likewise, several dynamical systems researchers who have built research programs by observing systems without central control mechanisms have acknowledged a role for cognitive elements in coordinated behavior. Examples include Pellecchia and Turvey (2001) who noted evidence that cognition affects the dynamics of a coordinated system, and Kelso (2003) who described cognitive event-related brain dynamics in coordination. Despite these theoretical developments, much work remains before any consensus on the components of handwriting theory will be achieved.

In the next sections, I will describe representative handwriting studies from information processing and dynamical systems theories. In most of the studies I will cite, participants drew ellipses, loops, straight disconnected lines, or spirals with an electronic pen on a digitizing tablet while seated at a table or desk. In some cases, test stimuli were presented on the writing paper, while in other studies, test stimuli were presented on a vertical computer screen placed at eye-level behind the writing surface.

Many studies varied the rates or sizes of the drawings and manipulated the visual trace, presenting delayed, distorted, or rotated output, or suppressing it altogether. Occasionally, studies involved writing nonsense words (Burton, Pick, Holmes, & Teulings, 1990; Longstaff & Heath, 1997, 1999), but few studies involved participants in performing meaningful writing tasks. Examples of research in functional handwriting skills have examined developmental change in writing skill in young children (Sage, Zesiger, & Garitte, 2009; Vilageliu & Kandel, 2009) and functional writing (words and numbers) by elderly individuals with mild cognitive impairments (Werner, Rosenblum, Bar-On, Heinick, & Korczyn, 2006). Additionally, forgery and deception studies are often realistic, requiring participants enter words and numbers on documents such as blank check forms or to write true and false declarations (Dick, Rogers, Found, 2000). Nevertheless, the majority of handwriting studies require participants to draw simplified written shapes without semantic content, such as loops, lines, and spirals. These simplified experimental tasks highlight the difficulties involved in studying multilevel psychomotor systems, and have prompted criticism that such studies do not represent natural handwriting. Although loops, lines, and spiral formations, are foundational to Western alphabets (Overvelde, 2009; Schmandt-Besserat, 1996), those forms do not incorporate linguistic elements of natural handwriting, and therefore, may not fully reflect the behavior they are designed to illuminate. Thus, important as the studies of simplified handwritten forms have been for understanding the motor processes underlying complex handwriting motions, their contributions to a complete understanding of handwriting are limited.

2.4. Information Processing Theory in Handwriting Research

According to the classic information processing perspective, handwriting is a system of numerical and alphabetical symbols depicted sequentially and interpreted by a reader (human or computer) through symbol-processing abilities (Brown, McDonald, Brown, & Carr, 1988). Originating in the brain, these processes are generated by a set of instructions, or motor programs, and operate in a top-down direction within a closed loop system. Corrections and guidance occur as errors are detected and corrected while feedforward mechanisms direct movements (Stelmach, Mullins, and Teulings, 1984; Teulings and Schomaker, 1993). Since motor programming has had a prominent role in handwriting research, I will first describe representative studies describing evidence for motor programming in handwriting, and then discuss the role of cognition and vision in handwriting according to the information processing perspective.

2.4.1. Motor Program Concepts in Handwriting Research

Van Galen & Teulings (1983) tested motor program theory by requiring participants to write the letter h repeatedly with 16 different parameters based on combinations of varied letter sizes, letter orientations, and writing direction. Twelve participants were instructed to write a series of six *h*'s (***hhhhhh***) with an electronic pen on a digitizing tablet as quickly and accurately as possible. Each trial consisted of the letter *h* repeated six times in each of 16 conditions based on a 2 (size) x4 (orientation) x2 (direction) factors: (i.e., size (0.33 or 1.00 cm), orientation (normal horizontal baseline, baseline tilted 90 degrees rightward, the baseline tilted 90 degrees leftward, or upside down), and direction (normal forward or reversed backward). Lines or small dots cued the participants to the requested condition and no online visual feedback was provided. The researchers measured the order in which strokes were written in each letter, along with reaction times, movement times, and peak velocities. They concluded that changes in movement time accompany changes in size and order of the strokes, and reflected stages of motor planning and initiation that were consistent with motor program theory. From these findings, Van Galen and Teulings (1983) inferred that handwriting was likewise governed by a three-stage motor program that involved "(a) retrieval of a motor program, (b) setting the parameters for that motor program, and (c) recruiting the appropriate motor units for starting the sequence of movements" (p. 21).

In a variation of the Stroop test for handwriting, Stelmach, Mullins, & Teulings (1984) explored the operation of motor programs investigating reaction time, considered to indicate the operation of motor programs (Klapp, 1996). In this study, they required 13 research participants to write a set of meaningless allographs such as hye as fast as possible after being prompted by a stimulus allograph flashed onto a computer screen. Participants were instructed to reproduce the shape of the stimulus allograph between a set of vertical lines. In half the trials, the participants reproduced the allograph in a 1.5 cm space; in the remaining trials, the participants reproduced the stimulus allograph in a 3 cm space. The researchers varied the size of the stimulus allograph so that the stimulus matched the drawing space in only 80% of the trials. In the remaining 20% of trials, the stimulus was considerably larger or smaller than the drawing space that participants were asked to fill. Researchers gave feedback after each trial to assist participants to match the size of their writing to the size of the drawing space.

The results showed that reaction times were much faster under conditions where the size of the allograph stimulus matched the size of the drawing space that participants were allowed. From this study, the researchers concluded that participants were able to successfully plan their movements when the stimulus matched the spatial features of the task, but that more processing time was needed when the stimulus did not match the required response. These findings also led them to conclude that relative timing in stroke sequences was an invariant feature of handwriting motor programs.

2.4.1.1. Motor Equivalence in Handwriting Research

In relation to handwriting, motor equivalence refers to the ability to achieve the same output after limited practice whether writing by dominant or non-dominant hands, the mouth, or the feet (Bernstein, 1967; Lashley, 1933; Meulenbroek et al., 1996). Motor equivalence has posed vexing problems for the motor program theory since it challenges the traditional motor programming notion that only one set of commands is responsible for attaining a specified motor goal. Since the same muscular contractions cannot be involved in writing with dominant and non-dominant hands, motor equivalence suggests that motor programming cannot specify a unique set of muscular commands, but refers instead to abstract processes of motor planning, preparation, and initiation.

Motor equivalence has been observed anecdotally in individuals who have lost the ability to write with their dominant hand but learned to write patterns with their non-dominant hand that resembled their original writing. Perhaps the most famous example of motor equivalence was shown by the similarity of handwriting by opposite hands written by Horatio Lord Nelson who lost his dominant right arm in a sea battle

(Perl, 1955). In 1993, Castiello and Stelmach explored this view on motor equivalence in a single case study involving a participant who had lost his dominant left hand and arm in an automobile accident. He had become proficient at writing with his nondominant right hand and, later, with a prosthetic myeoelectric arm and hand on his still-dominant left limb. Researchers included an intact left-hand dominant participant as a control for comparison. In the study, the participants wrote their first names in their typical cursive style of writing and the equation X + Y = 2 on a digitizing tablet, using three effectors: their dominant left limb, dominant left elbow, and non-dominant right hand. The disabled participant wrote with the myoelectric prosthesis attached to his dominant left arm, and was highly practiced in the use of each of the three effectors. The control participant was highly practiced in writing only with the dominant left hand and was unfamiliar with writing with the left elbow and non-dominant right hand. Kinematic analyses of the patterns of movement by the disabled and intact research participants showed that highly practiced writing, regardless of effector, exhibited similar kinematic features. Thus, the kinematics of the disabled participant, who was skilled in writing with all three effectors, showed few differences across effectors. However, the kinematics of the control participant showed significant differences between the practiced and unpracticed writing movements. The researchers concluded that motor equivalence could be explained by the operation of a generalized motor program that is activated by practice.

2.4.2. Cognition in Information Processing Theories of Handwriting

Van Gemmert and Van Galen (1997) was among the first studies to explore cognition as it relates to handwriting, which produced a theory of neuromotor stress in which “cognitive and physical stressors enhance the level of neuromotor noise in the information-processing system [which] facilitates easy tasks but disrupts complex tasks” (Van Gemmert & Van Galen, 1997, p. 1299). In their study, they varied a series of experiments with tasks involving number writing or graphical aiming (i.e., drawing a line to a target) in each of eight conditions based on a 2x2x2 design: (a) task difficulty (low or high), (b) cognitive stress (low or high), and (c) physiological stress (low or high). Participants wrote on a digitizing tablet with their recorded traces (numbers or aiming trajectories) displayed on a computer monitor facing them. In the lower-difficulty number-writing task, participants were required to write many single digits as quickly as possible on a digitizing tablet; the higher-difficulty number-writing task required participants to write double-digit numbers as quickly as possible. These tasks were paired with two levels of cognitive stress and two levels of physiological stress. The cognitive stress conditions involved repeating the same number while writing (low cognitive stress) or stating the results of serial subtraction while writing (high cognitive stress). At the same time, participants were subjected to two conditions of physiological stress consisting of a 55 dB(A) auditory tone (low physiological stress) or a 95 dB(A) auditory tone (high physiological stress). The results of these trials showed that writers responded to different levels of cognitive or physiological stress by increasing stiffness of the effectors producing the movement, exploiting the friction properties of the writing surface, or changing the characteristics of the task.

In the graphical aiming experiment in the study, Van Gemmert and Van Galen (1997) varied the experiment by requiring participants to perform an aiming task under the same conditions of cognitive and physiological stress as the number-writing task. Participants were asked to draw lines of different lengths (moderate difficulty, 2.5 cm; or higher difficulty, 5 cm) on paper placed on a digitizing tablet, beginning at different starting points directed toward circles placed around the paper. As in the previous experiment, the traces of the participants' drawings were presented on a computer screen facing them. For the graphical aiming task, the researchers found that the participants responded to the increased cognitive stress through increased axial pressure.

The results of the number writing and graphical aiming tasks led these researchers to observe that "increased processing demands [during motor performance]…lead to increased levels of …decreased signal-to-noise ratios in the system" (p. 1300). They concluded that "performance control is as much the outcome of the constraints of the body as the result of an optimized balancing between perceived accuracy demands" (p. 1311). These researchers proposed that a comprehensive theory of handwriting performance should incorporate dynamic and biomechanical factors along with information processing mechanisms to explain consistency and variability in handwriting.

Information processing models of handwriting have been applied to computing as well as psychological and clinical issues. For example, Plamondon and Srihari (2000) reported that models of human decision-making and judgment, inferred from studies of handwriting dynamics, has been used to develop computer algorithms for both handwriting recognition and signature verification. The study of handwriting dynamics has also been useful in gaining insights into human-computer interface. From observations of pen and associated eye movements during writing tasks, Caporossi, Alamargot, and Chesnet (2004) built mathematical models that may be used to enhance human-computer communication.

In addition to the varied psychological, clinical, and computing applications of cognition, information processing models of handwriting have furthered understanding of language learning. Growing numbers of research reports have linked handwriting with improved quality of written language by students from elementary grades through high school (Berninger et al., 2009; Connelly, 2009; Kandel, 2009; Medwell, Strand, & Wray, 2008; Shanahan, 2008). These studies found that students who developed automaticity in writing produced higher quality writing, characterized by more complex thought patterns when compared to the work of students who lacked such practice. As a result, these researchers have joined a growing number of educators who urge a return to handwriting practice as a method for teaching and refining composition skills and critical thinking abilities.

2.4.3. The Role of Vision in Information Processing Views of Handwriting

The role of vision in guiding limb movements, broadly defined, has been studied extensively (see Ketcham, Dounskaia, & Stelmach, 2006, for a review). According to the information processing model, visual operations are used with cognitive judgments to assist in preplanning movements and guide the motor system through feedback and feedforward mechanisms (Adams, 1971; Adams & Goetz, 1973; Seidler, Noll & Thiers, 2004; Schmidt & Wulf, 1997; Winstein, Pohl, & Lewthwaite, 1994). Other researchers have explained that information processing models account for improved spatial accuracy (Elliott, Garson, Goodman, & Chua, 1991) and motor learning (Wishart, Lee, Cunningham, and Murdoch, 2002) with vision as well.

Beyond the general view that vision benefits movement, the specific role of vision for highly skilled aiming and writing tasks is less clear. For example, Miall (1996) and Proteau and colleagues maintain that spatial and timing accuracy is enhanced with vision in stylus-aiming skills, even for actions in which movers are well-practiced (Proteau & Cournoyer, 1990; Proteau, & Marteniuk, 1993; Proteau, Marteniuk, Girouard, & Dugas, 1987; and Proteau, Tremblay, & Dejaeger, 1998). These researchers proposed that visual feedback in well-practiced skills allows the subject to exploit benefits of vision not detected at earlier stages of motor skill learning.

Some proponents of information processing theory and dynamical systems theory have predicted opposite effects from concurrent visual feedback. For example, Brown & Donnenwirth (1990) claimed that vision in hand motions increases task difficulty by requiring concurrent monitoring of motor and cognitive components of aiming tasks. Several researchers have proposed that concurrent augmented visual feedback increases the amount of information to be processed by the perceptual-motor system, increasing task difficulty (Berryhill, Kveraga, Boucher, & Hughes, 2004; Kording & Wolpert, 2004; Marquardt, Gentz, and Mai, 1999; and Swinnen, Schmidt, Nicholson, and Shapiro, 1991). Their findings are consistent with the features of a dual task paradigm (Welford, 1973) which hypothesizes that task demands compete for cognitive resources.

In contrast to the controversies regarding the role of visual monitoring of movement in aiming and tracking studies, the handwriting research has consistently found few benefits of ongoing vision during skilled writing. Vision appears to benefit writers by assisting with spatial alignment and placement of written or drawn objects (Graham and Weintraub (1996), Rijntjes, Dettmers, Büchel, Kiebel, Frackowiak, and Weiller (1999), but ongoing visual monitoring of hand motions during writing adds little additional advantage (Smyth & Silvers, 1987). Several researchers have observed that characteristic features of the handwriting of separate individuals persist due to the joint influence of cognitive and biomechanical factors (Meulenbroek, Thomassen, van Lieshout, & Swinnen, 1998). Adherents to motor programming theories explain that continuous visual guidance is unnecessary because handwriting processes are initiated and controlled by central programming mechanisms (Van Galen, Smyth, Meulenbroek, & Hylkema, 1989). Other researchers explain the limited benefit of continuous visual monitoring of writing by suggesting that visual feedback is not processed quickly enough to be useful in most skilled writing. For example, Graham and Weintraub (1996), commenting on the research done by Teulings and Schomaker (1993), noted that "The role of vision in guiding the moment-to-moment movements involved in handwriting [is limited] because visual feedback (as well as kinaesthetic) is probably too slow to be effective at the speed adults usually write" (p. 16). In an fMRI study of writing, Teulings and Schomaker (1999) reported that activity in visual areas of the brain increased by only small amounts during a set of handwriting and footwriting tasks in normal adults, presumably because feedback from the visual system is slower than brain activity in the motor centers.

Although studies agree that vision is not needed to produce writing, they have produced mixed results for determining the effect of continuous visual monitoring on speed during handwriting. Marquart, Gentz, and Mai (1996) determined that visual attention to handwriting forces handwriting movements to become slower. Invoking the information processing model, they explain that vision "disables the generation of automated [handwriting] movements …which are typically in the range of 4-6 Hz…caus[ing] the movements to slow down" to 1-2 Hz, the limit that the human eye can successfully track (p. 96). Kao, Shek, and Lee (1983), found similar slowing in movement times with increased task complexity).

Results from Chua and Elliot (1997), in which participants moved a cursor toward a target under vision and no-vision conditions, found similar results in a study of visual-mediated closed loop control of aiming movements. These researchers instructed ten right-handed participants to move from a home position to a target 130mm away on a digitizing tablet placed horizontally in front of them. Both the starting position and the targets were presented on a computer screen placed vertically at eye level behind the digitizing tablet. Participants moved as quickly and accurately as possible from the starting position to the target in each of six conditions that varied combinations of concurrent visual feedback of the cursor movement (present or absent) and the target size (5mm, 10mm, and 20mm). Their results showed significant increases in movement time (i.e., slower speeds) but improved consistency of movement with the visual trace of the cursor movement.

In contrast to these findings, several research studies have shown that vision enables faster movement times in handwriting (Van Doorn & Keuss, 1992; Van Galen, Smyth, Meulenbroek, & Hylkema, 1989). For example, Van Doorn and Keuss (1992) required twelve adult participants to write meaningless sequences of letters with and without vision. No significant differences in letter shapes were found between the vision and non-vision conditions, but reaction times and movement times were longer in the non-vision conditions. These researchers concluded that removal of vision impeded movement times and lengthened reaction times because the planning stages of the underlying motor programs were fostered with vision, and impeded without it.

The range of views regarding the effects of vision on handwriting speed and tracking tasks illustrate the challenges involved in the scientific study of graphonomics. Differences among the studies presented in this section reflect varied research questions and methods (Olive, 2004). Thus, the various positions presented by these authors may not be in disagreement, but instead form complementary views on multilevel processes in handwriting, aiming, tracking, and drawing activities.

2.5 Dynamical Systems Concepts and Handwriting Research

Handwriting and drawing are activities that share some of the same dynamical characteristics and origins that have been found in many other motor behaviors, from finger tapping to hip-ankle coordination in stance (Stoffregen, personal communication, April 26, 2007). Although a few early studies of dynamical systems (Denier van der Gon and Thuring, 1965; Vredenbregt & Koster, 1971) used handwriting to observe oscillatory behavior, modern studies of handwriting dynamics are few. Those that have incorporated dynamical systems concepts have addressed coordination dynamics, attractor states, oscillatory (i.e., rhythmic) motions, and frequency coupling. For example, the concept of coordinative structures has been used to explain the stability of oscillatory movements and coupling functions between different limb-limb and perceptual system–motor system components (Kelso & Schöner, 1988; Meulenbroek, R.G.J., Thomassen, A.J.W.M., Schillings, J.J., & Rosenbaum, D.A. (1996); Sternad, Turvey, & Schmidt, 1992). The research group of Sallagoïty, Athènes, Zanone, and Albaret (2002, 2003, 2004) has been instrumental in demonstrating coordination dynamics in handwriting.

In this section, I will first describe representative studies describing complex nonlinear dynamics in handwriting research. Then I will describe the applications of coordinative structures and attractor states to contemporary handwriting research, followed by a discussion of the possible roles of entrainment and vision for inducing timing and spatial accuracy in tracking tasks.

2.5.1. Early Research on Dynamical Principles in Handwriting

In the 1970’s, researchers in movement sciences were seeking new models to address challenges posed by ecological psychology (Gibson, 1972) to traditional information-processing explanation of perception and action. At the same time, the work of Nikolai Bernstein (1967) who conceptualized movement as the product of a system, a “natural phenomenon, encompassing the interactions between the brain, the movement system, and the natural and cultural environments” (Bongaardt & Meijer, 2000, p. 68), emerged as a partial response to these issues. Bernstein’s lengthy experiences observing movements led him to conclude that pre-programmed motor commands could not produce all the skilled flexible and stable features of coordinated activity required in natural, unpredictable environments, an observation described as Bernstein’s degree of freedom problem. In 1977, Turvey wrote *Preliminaries to a Theory of Action with Respect to Vision*, creating a new framework that applied concepts from Bernstein (1969) to address the criticisms of the information processing approach to perception and action posed by ecological psychologists. In this work, Turvey proposed concepts from the field of mathematical dynamical systems to challenge the primacy of information processing models in motor control. A later work (Turvey & Carello, 1981, 1986) linked ecological psychology with the observations in Bernstein’s studies and with dynamical systems. Researchers (e.g., Kugler, Kelso, & Turvey, 1980) thus began testing hand motions and handwriting for evidence of dynamical systems. The seminal paper that launched applications of complex dynamical systems in handwriting research was written by Hollerbach (1981), a computer scientist and mechanical engineer from MIT. Hollerbach’s study involved the use of a digitizing tablet to measure hand position, an accelerometer, and a frictionless X-Y sliding rail system used to obtain continuous measures of velocity and acceleration. In this study, participants were required to write their signature and connected strings of lower case cursive letters such as *e*s (i.e., ***eeee***), *e*s alternating with *l*s (i.e., ***elelelel*** or ***eelleell)***, *u*s (i.e., ***uuuu***), *m*s (i.e., ***mmmm***), and other letters; and short words such as *hell* and *elye*). By mathematically analyzing the velocity and acceleration profiles of the writers' letters and hand movements, he demonstrated that handwriting could be modeled as a system of "coupled oscillations in horizontal and vertical directions …superimposed on a rightward constant velocity horizontal sweep" (p. 139). Although Hollerbach endorsed motor program concepts, his work revealed handwriting to be a dynamical system by demonstrating that letters could be formed through manipulating the dynamics of oscillators.

*2.5.2. Principles of C*oordination Dynamics Applied to Handwriting

Bernstein's solution to the degree of freedom problem in coordinated movement was the *synergy*, (more commonly termed *coordinative structure*), defined as a set of motor units flexibly assembled to achieve a movement goal, operating as an interdependent system and constrained by the demands of the task. Turvey and colleagues adopted this notion, arguing that coordination emerges as a self-organized property of interacting coordinative structures without direction from a central controller, such as a motor script or program (Tuller, Turvey, & Fitch, 1982). This theory criticized the view that movements are primarily top-down processes, dependent on preplanning based on mental representations and stored motor commands, and initiated a breach in movement research between information processing and dynamical systems approaches that has continued.

Haken, Kelso, and Bunz (1985) were among the earliest researchers to propose that coordinative structures were assembled by nonlinear coupling of time dependent movement components, cooperating through the mechanism of relative phase. In an early study to test for coordinative structures in handwriting, Van Emmerik and Newell (1989) required research participants to write their names and sequences of connected cursive *e*s (i.e., ***eeeeee***) with their non-dominant hands. The researchers observed that the writers exhibited freezing of degrees of freedom between proximal and distal joints while performing this unfamiliar task, as is typical of novice learners of other forms of coordinated activity. They concluded that the processes of freezing, releasing, and exploiting degrees of freedom were a typical sequence in motor learning, and determined that motor learning leads to successful movement outcomes by incorporating knowledge of agent-task affordances and biomechanical constraints.

While many researchers have not endorsed the dynamical systems framework as the sole explanation for movements in handwriting, they agree that self-organized coupled oscillations within the psychomotor system exert powerful influences over motor aspects of skilled handwriting. For example, Dounskaia, Van Gemmert, and Stelmach (2000), in a study of circle and line drawing, varied speeds and combinations of wrist and finger flexions to observe coordination. They concluded that intrinsic biomechanical constraints in low relative phase relationships, rather than cognitive processes, dominated the production of cursive shapes.

Of particular relevance to the present study, principles of dynamical systems have been supported in studies of tracking motions. For example, Liao and Jagacinski (2000), who conceptualized rhythmic tracking tasks a "coupling between oscillatory limb movements and an oscillatory external visual signal" (p. 362), designed a study to observe the relationship between hand motions and a rhythmically moving target. In this study, they paired two visual information conditions (real-time movement of the cursor vs. real-time graphics depicting tracking errors) with two motor control conditions (accurate scaling and orientation of the real-time visual trace vs. accurate scaling but displacing the visual trace at a 90° from the actual tracking trajectory). The researchers found that participants had the fewest tracking errors when they maintained 0° or 180° relative phase with the oscillating target which was successful only when the visual information on the screen reflected the actual cursor trajectory without the 90° displacement rather than the graphical depiction of the necessary correction to reduce movement error. However, the optimal relative phase relationships broke down as target speed increased, decoupling the agent's motions from the target. Liao & Jagacinski (2000) concluded that the coupling between the oscillating limb of the participant and the oscillating target was an effective coordination strategy at speeds within agent's perception-action capability.

2.5.2.1. Attractor States.

The hallmark of dynamical systems is the attractor state, a stable set of self-sustaining patterns which the system gravitates toward and maintains over time, unless the system is perturbed (Strogatz, 1994; Williams, 1997). This feature of dynamical systems has been used to explain that *"*although an enormous range of [behavioral] patterns is theoretically possible, the system actually displays only one or a very limited subset of them….The system "settles into" or "prefers" only a few modes of behavior" (*Thelen & Smith, 2006). In other words,* behaviors tend to drift toward a nearby stable pattern, termed attractor basins. As a result, individuals engaged in a particular task under specified conditions, such as handwriting or gait, produce behavioral patterns that are easily recognizable. In coordinated activity, stable attractor states emerge from the physical properties and forces produced by the moving organism when coupled to a constant stream of behaviorally relevant information (Ouillier, Bardy, & Bootsma, 2001).

The stability of attractor states are shown by their resistance to the effects of small perturbations applied to the system and a rapid return to stability when such a perturbation occurs (Kelso and Schöner, 1988). Thus, responses to perturbations reveal much about the relationship between the motor system and the influence of vision, an important source of behaviorally relevant information, on the production of well-practiced movements (Marquardt, Gentz, & Mai, 1999).

Principles of attractor states applied to handwriting. Athènes and colleagues (2003) applied contemporary technology to reevaluate Hollerbach (1981) who proposed that cursive letters in English were formed by *x-y* oscillations translated in a rightward direction over the page, and extended their inquiry to include American English, modern European, Arabic, Chinese, and Hebrew writing systems. They determined that *x-y* oscillations comprised the majority of alphabetic forms in each of these languages, concluding that the alphabets used by the vast majority of the world's population can be rendered as oscillations of x-y coordinative systems. Since oscillatory movements in biological systems naturally converge toward and remain within attractor states, often within a few cycles (Strogatz, 1994; Williams, 1997), Athènes and colleagues were motivated to investigate elliptical handwriting motions for evidence of attractors.

In studies published in 2002, 2003, and 2004, Sallagoïty, Athènes, Zanone, and Albaret instructed six participants to draw circles and ellipses of different phase relationships and orientations (slants) in 15 degree increments and relative amplitudes, as shown in Figure 1a and Figure 1b.

Figure 1a. Ellipses oriented at different angles

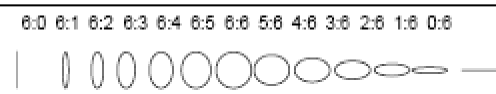

Figure 1b. Ellipses oriented at different relative phase

Figure 1. Ellipses used by Sallagoïty, Athènes, Zanone, and Albaret (2002, p. 137) to determine pen motion attractor states during elliptical drawing.

Figure 1a introduces a sequence of forms beginning at 0 relative phase starting with a straight line oriented at 45°, followed by a series of ellipses with the major axis oriented at 45°, expanding across an orthogonal minor (i.e., vertical) axis. Six subsequent ellipses expanded the ratio of major to minor axis in 1/6 increments (6:1, 6:2, 6:3, etc.) until reaching a 6:6 ratio (a circle) depicting an equal relationship between maximum ellipse height and width. After the circle, the sequence continues with an additional set of ellipses, with the major axis oriented leftward 135°. Each successive ellipse was formed by contracting the width across the minor axis in 1/6 increments. A straight line oriented at 135° completed the series.

These researchers also analyzed the effect of varying the amplitude of ellipses in relation to height as shown in Figure 1b, where the horizontal and vertical axes were fixed orthogonally at 0° and 90°, respectively. This series of ellipses shows successively changing ratios between major (horizontal) and minor (vertical) axes from 6:0 to 6:6, as in the previous set of ellipses, depicted by Figure 1a.

Of these forms, four stable attractors emerged, characterized by precision, velocity, reduced error, and reduced variability. The four forms that met the criteria for an attractor consisted of the following shapes: (a) straight vertical lines (6:0 phase relationship), (b) horizontal lines (0:6 phase relationship), and (c) two intermediate ellipses: one elliptical form with a phase relationship of 6:3 (the ellipse being twice as tall as it is wide) and one elliptical form with a phase relationship of 3:6 (the ellipse being twice as wide as it is tall), which were mirror images of each other.

As a result of these experiments, Athènes, Sallagoïty, and colleagues (2003, 2004) found that the most natural handwriting configurations for an individual writer fall into a small set of patterns of coordinated hand movements. They concluded that these preferred movements reflect attractor states emerging from non-linear coupling of the hand with the dynamics of the task, which in this case are the two *x*-*y* dynamic oscillators comprising the elliptical shapes. Thus, attractor states may provide an explanation for stable handwriting features for a particular writer, consisting of a narrow range of values of handwriting parameters (force, amplitude, velocity, phase relationships, and other handwriting components).

2.5.2.2. Synchrony as a Characteristic of Coordinated Movement.

Many studies reveal that individuals coordinate the amplitude and timing of their movement rhythms with events in the environment that they see or hear, including the movements of other people. The capacity of subjects to synchronize movements with external stimuli is a well-documented observation about motor behavior (Patel, 2006; Repp, 2001, 2002, 2005; Strogatz, 2003). "[F]or reasons we don't yet understand, the tendency to synchronize is one of the most pervasive drives in the universe, extending from atoms to animals, from people to planets" (Strogatz, 2003, p. 14). Examples of behaviors found to be coordinated with external events include a wide range of behaviors, such as posture (Carson, 2006), limb movements (Brown, Martinez, & Parsons, 2006; Richardson, Schmidt, & Kay, 2007), musical performance, (Trainor, 2007), conversation (McGrava & Warner, 2003; Shockley, Santana, & Fowler, 2003), interpersonal coordination (Fowler, Richardson, Marsh, & Shockley, 2008; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; Schmidt & Turvey, 1994) and infant-parent bonding (Feldman, 2007).

In particular, studies have shown humans to be highly sensitive to temporal features of moving stimuli, resulting in frequency coupling across joints and limbs within the body and with moving stimuli in the environment. As a result, individuals detect changing timing patterns and adjust their movements accordingly, often without any conscious awareness of doing so. Studies of coordinated activities that require close frequency coupling for successful performance include behaviors as diverse as juggling (Beek, 1986; Tlili, Mottet, Dupuy, & Pavis, 2004), tapping and drumming (Clayton, Sager, & Will, 2004; Jones & Boltz, 1989; Repp, 2005), speech (Saltzman & Byrd, 2000), finger movements (Kelso, Schöner, Scholz, & Haken, 1987), and bimanual coordination (Bingham, 2004). In these studies, timing in coordinated movement is attributed to frequency coupling, defined in terms of relative phase paradigms which describe the spatial relationships between two oscillating systems (Balasubramaniam, 2006; Burgess-Limerick, Abernethy, Neal, 1993; Kelso, 1997; also see Richardson, Marsh, & Schmidt, 2005, and Volman & Geuze, 2000, for a review).

The adaptive benefit of coupling phenomena in human movement is revealed by adverse effects that result when individuals do not synchronize the timing of their movements with events in the environment. Although problems arising from unsynchronized motions may be clear in some situations, as in the case of auto accidents (Consiglio, Driscoll, Witte, & Berg, 2003), others may not be obvious to the individuals involved, but nonetheless have significant effects. For example, Finkel, Campbell, Brunell, Dalton, Scarbeck, and Chartrand (2006) found that interpersonal synchrony of subtle movements was higher in congenial workplaces, and lower where there were higher rates of cognitive and emotional stress, resulting in lower productivity. Observations such as these suggest that temporal coordination in relation to dynamic or static environmental features is fundamental to successful adaptation.

Researchers have searched for timing mechanisms that enable the process of synchronicity across components of the body, among individuals, and between individuals and external oscillating objects, but results have been elusive. Some researchers have invoked central timing mechanisms in visuomotor synchronization (Kurgansky, 2008) or the concept of biological time keeping mechanisms such as internal clocks (Ivry & Richardson, 2002) to explain timing in rhythmic behaviors such as cardiovascular or locomotor rhythms (Hill, Adams, Parker, & Rochester, 1988; Niizeki & Miyamoto, 1998) and bimanual drawing motions (Zelaznik, Baria, Bloom, Dolansky, Justice, Patterson, Whetter, Spencer, & Ivry, 2005).

Conversely, dynamical system models of behavior challenge the concept of an intrinsic timekeeping mechanism. According to the dynamical systems theory, timing is a feature of coordinated motion, emerging as perceptual-action systems are coupled through direct visual, auditory, and kinesthetic processes occurring in real time (i.e., "online") (Gibson, 1979; Turvey, 1977). To ecological psychologists and dynamical systems theorists, coupling between perception and action is direct, informational, and reciprocal – we perceive in order to move, but we move in order to perceive (Shaw & Schockley, 2003), without intervening mental processing or mental representations.

Thus, instead of a cellular structure or mental construct functioning as a biological clock to foster temporal coupling, dynamicists suggest instead that timing emerges from attractor states and processes driven by visual or auditory signals (Fink, Foo, Jirsa, & Kelso 2000; Phillips-Silver & Trainor, 2007; Repp, 2001; Robertson, Zelaznik, Lantero, Bojczyk, Spencer, Doffin, & Schneidt, 1999; Schöner, 1991; Staude, Dengler, & Wolf, 2002; and Wimmers, Beek, & Van Wieringen, 1992). Wann and Nimmo-Smith (1990) contend that “there is little experimental evidence that directly supports time as an explicit parameter in the process of specifying a motor response that cannot be equally well accounted for by control models based upon oscillatory systems and point attractors" (pp. 115-116).

Entrainment as a special case of synchronous behavior. Entrainment is a special case of synchronized behavior pertinent to tracking tasks. Over a century ago, researchers observed that temporal patterns in movement exhibit drift, typically slowing across a movement sequence unless influenced by a rhythmic auditory or visual signal to foster temporal coupling (Dunlap & Wells, 1910; Stetson, 1905). By introducing an oscillator into the system, the process of drift is reversed, resulting in frequency coupling, an entrained response.

Entrainment, therefore, is the convergence of two or more interacting oscillating processes pulled toward a common periodicity in amplitude, frequency, or both (Clayton, Sager, & Will, 2004). Weltman, Koepke, and Selcher (2000) note that “spatiotemporal periodic pulling [is] a specific but universal phenomenon associated with driven, nonlinear, spatiotemporal systems” (p. 2773). Thus, entrainment has been a robust finding in many movement studies, with broad implications for the motor and cognitive aspects of coordination (Bernstein, 1967; Kelso, 1995; Kelso, DelColle, & Schöner, 1990; Port, Tajima, & Cummins, 1996; Treffner and Turvey, 1993). In many cases, the temporal relationships between the oscillating systems begin out of phase, but shift toward the driving stimulus, often outside the mover's awareness (Jones & Boltz 1989; Weaver, Lobkis, & Yamilov, 2007), until the system converges toward the driver in simple ratios such as 1:1, 1:2, or 1:3 (Port, Tajima, & Cummins, 1996). Richardson, Lopresti-Goodman, Mancini, Kay, and Schmidt (2008) demonstrated that rhythmic coordination – both intrapersonal and with external stimuli (including other actors) - are "both constrained by the self-organizing entrainment process of coupled oscillators" in the systems (p. 340). Further, they observed that coupling is most effective when the oscillation of the external stimulus is close to a preferred movement frequency of the agent (Lopresti-Goodman, Richardson, Silva, & Schmidt, 2008).

One major indicator of entrainment is the difficulty found in resisting its effects. In their study on posture, Stoffregen, Hove, Schmit, and Bardy (2006) found that individuals who become aware that their posture is coordinated with some external stimulus can resist it to some degree, but only with intentional effort. The mechanisms of resisting entrainment, however, are not fully understood.

Entrainment between hand motions and events in the environment. Hand motion studies were among the first to be studied for responsiveness to entrainment (Kay, Saltzman, Kelso, and Schöner, 1987; Swinnen, Dounskaia, Walter, Serrien, 1997).

Despite extensive studies on entrained finger movements and manipulations of pendula, entrainment in handwriting and drawing activities has received little attention, perhaps because any such study requires a simultaneous investigation of cognitive, perceptual, and motor processes (Repp, 2005). As handwriting and drawing rely on biomechanical-neuromuscular factors and cognitive mechanisms to match a mental representation of a particular shape, research on entrainment in handwriting and drawing is a complex study that requires teasing out the relative contributions of physical and cognitive components of writing.

2.5.2.3. Dynamical Systems and Visually Guided Hand Motion

Many researchers have found vision to foster coordinated timing in movement (Keogh, Morrison, & Barrett, 2004; Turvey, Shaw, & Mace, 1978), proposing that vision promotes coupling between the stimulus and relevant components within the motor system. However, auditory or multimodal rhythm devices, such as a metronome, have been found to be more effective at maintaining synchrony between movement and inanimate oscillating targets than vision alone (Chen, Repp, & Petel, 2005; Kudo, Park, Kay, and Turvey (2006). Of particular relevance to the present study, Ceux, Buekers, and Montagne (2003) found entrainment between agents and inanimate signals to depend on the quality of the perceptual motor coupling in a tracking task. In their study, they instructed participants to synchronize tracking movements to a target consisting of an oscillating light. They manipulated the spatial-temporal features of the task and the nature of the visual information by providing online visual traces of the movement in some trials, while suppressing feedback in others. Their findings demonstrated that visual feedback mediated temporal coupling to significant degree, since removing visual feedback disrupted both spatial and temporal coupling. Further, they found that entrainment did not occur continuously along the task trajectory, but at anchor points, and most notably at reversals along the trajectory. Further, they note that agents may synchronize differently to animate and inanimate movements. In a more recent study, Gill (2007) concurred with the finding of intermittent entrainment, finding that attention directed toward the end-points of alternating visual stimuli reduced variability and increased stability in the visuo-motor response.

In addition to the effects of vision on temporal coupling between hand motions and an oscillating stimulus, several dynamical systems theorists have proposed that concurrent visual monitoring of hand motions during writing may influence the emergence of attractor states within the actor-task system (Athènes, Sallagoïty, Zanone, & Albaret, 2004; Chen, Weng, et al., 2008), producing motions that are constrained by properties of the task-agent system, and not driven solely by the task demands. Overvelde (personal communication, September 16, 2009) has observed that novice writers practicing hand movements that are tightly constrained to the task, as when tracing the grooves of letters that are carved into a wooden template, exhibit considerably less motor learning of letter shapes than those movements that are less restricted, as when writing on paper, presumably because an attractor state cannot emerge with such constraints.

2.5.2.4. Dynamical Systems and Cognition in Coordinated Hand Motion

Information processing theorists criticize the dynamical systems approach to handwriting by noting that the cognitive aspects of meaningful handwriting are ignored. Admittedly, the majority of research on dynamical systems in hand motions has addressed only the biomechanical, neuromuscular, and perceptual aspects of hand movements. Pellecchia, Shockley, and Turvey (2005) attempted to address this objection by conducting one of the few studies to incorporate cognition in a study of hand motions by investigating the effect of added cognitive tasks, such as counting-backwards, on various attractor states found in rhythmic hand motions. They concluded that concurrent cognitive tasks shift the properties of the preferred hand motions (i.e., the attractor) and makes its boundaries (i.e., the attractor basin) noisier, but does not interfere with the ability of the system to detect minor perturbations. Their results demonstrate that attractor states are sensitive to cognitive activity, and are not defined solely by physical processes. However, the scarcity of research in this area suggests that the influence of cognition on coordination is an open field of inquiry. A fuller understanding of handwriting dynamics requires more scholarly attention directed toward the cognitive influences on temporal coupling and attractors in handwriting.

2.6 Literature Sources for the Predictor (Independent) Variables

2.6.1 Shape, Amplitude, and Slant of the Loops Used as Target Stimuli

The tasks in the current study involve drawing connected twelve loops in repeated trials, some at a constant tempo, and others at changing rates. I chose loops for the experimental task in this study because the movements invoked when drawing loops are fundamental to the development of script (Overvelde, personal communication, September 16, 2009; Schmandt-Besserat, 2007). Further, loops have been commonly used in previous handwriting research and have been found to be compatible with handwriting coordination dynamics (Meulenbroek, Thomassen, van Lieshout and Swinnen, 1998; Sallagoïty et al., 2003; Thomassen & Meulenbroek, 1998). Meulenbroek et al. (1998) found loops to provide stable coordination patterns between wrist excursions and pen-tip displacements regardless of sample length or placement of the loop sequence on the paper. Athènes et al. (2004) and Sallagoïty et al. (2003) found that one of the most stable handwriting configuration was an elliptical shape drawn with an amplitude ratio of 6:3 between the major and minor axes (i.e., the ellipses were twice as high as they were wide at the widest point) and an orientation of 45°.

Using an experimental target associated with a known attractor state has been advised by Case, Tuller, and Kelso (2003) who noted that tasks compatible with stable patterns will engage coordinative structures, and reduce training time, motor stress, and cognitive stress during experimental tasks (Case, Tuller, Kelso, 2003, n.p.).

Thus, in this study, the shape of the interior area of the closed portion of the loops designed for the experimental task in the current study was based on the attractor features demonstrated by Athènes et al. (2004) and Sallagoïty et al. (2003) in the aforementioned studies.

2.6.2. Use of an Animated Target over a Static Template

In addition to selecting target shapes, sizes, and orientations to correspond with known attractor states, questions about the presentation of the target as a static or dynamic feature of the experimental task required consideration. A moving target was selected on the basis of Kaiser, Proffitt, Whelan, and Hecht (1992), who stated that an animated target improves accuracy in low dimensional tasks and aids subjects in predicting with greater accuracy the course of the moving trajectory.

2.6.3. Rates of Motion for the Animated Target in the Changing Rate Conditions

A changing rate for the target movement was included to investigate the effects of temporal perturbation on movement responses. In previous studies, rate changes were prompted by verbal instructions to write as fast as possible (Brown & Donnenwirth, 1990, Smits-Engelsman, Niemeijer, & van Galen, 2001) or to follow a smoothly changing rate of the stimulus in a pursuit rotor tasks (Morrens, Wezenberg, Verkes, Hulstijn, Ruigt, Sabbe, 2007).

Adding temporal perturbations to the pace of the visual stimuli in the present study is a novel approach in a drawing-tracking task. In order to introduce a perturbation in the system sufficient to force a change in the coordination dynamics of the response, the rate of the target animation abruptly changed at each of four successive loops, then resumed the constant rate. The rates for this manipulation were suggested by Daniel Kaplan (Kaplan, & Glass, 1995), Professor of Mathematics and Computer Science at Macalester College, (personal communication, March 15, 2006) and successfully led to sufficient variability in the timing of the drawing response. Repp (2002) has described that the tracking response of individuals to perturbations in the oscillating stimulus as a rich arena for future research.

2.6.4. Use of Concurrent Visual Feedback

Through advances in technology, researchers have been able to use concurrent augmented visual feedback, delivered through an external source such as a computer display (Van Dijk, 2006). This technology provides immediate feedback on the results of movement in the form of visual traces of movements, graphs, and visual displays.

Todorov, Shadmehr, and Bizzi (1997) found that augmented concurrent visual feedback fostered superior learning of a complex timing skills task than terminal or concurrent verbal feedback, but note that other research on augmented feedback has had mixed results.

2.7. Literature Sources for the Response (Outcome) Variables

2.7.1. Timing Errors: Timing Differences between the Target and the Participant

In this study, timing errors were defined as the time differences between the target and participants for reaching specific landmarks of the sample, i.e., the peaks of the twelve loops that comprised the sample. Calculating the timing differences at specific points on a cycle is a common method of measuring any timing error in movement studies (Newell, Carlton, & Seonjim, 1994).

2.7.2. Spatial Errors: Root Mean Square Error (RMSE)

Root mean square error (RMSE) is a common index of spatial accuracy and has been found to be reliable when comparing performance with and without visual feedback (Guadagnoli & Kohl, 2001). For example, it has been effective in producing accurate measures of feedback effects, i.e., knowledge of results in a series of studies (Guadagnoli, Dornier, & Tandy, 1996; Park, Shea, & Wright, 2000).

In this study, the method for calculating RMSE is an adaptation of a method used by Giraudo and Pailhous (1999). My goal was to apply a reliable measure of accuracy in reproducing a shape without penalizing for displacement errors, since participants were expected to make errors when reproducing the study shapes without visual feedback to guide placement. Although the discrepancy measure developed by Giraudo and Pailhous (1999) was designed for discrete strokes, it was adapted in this study to accommodate curved trajectories. This method is used for determining how closely the Euclidean distances between the endpoints within the target matched the Euclidean distances between endpoints within the sample without regard for the distance between a target loop and a sample loop by superimposed a target loop on each drawn loop to calculate spatial errors. This method is advantageous because it measures differences in shape alone, and does not penalize writers whose writing becomes misaligned on the digitizing tablet, a common occurrence when participants cannot see their own drawing.

2.7.3. Undulations (“Wiggle”) in the Writing Line

Fine, irregular undulations in the writing line were defined in this study as “wiggles”, and have also been described as wrinkliness (Chen, Cha, Chee, & Tappert, 2003), and dysfluency (Caligiuri, Teulings, Filoteo, Song, & Lohr, 2006). In comprehensive studies across adult age groups, wiggle has been found in the handwriting of normal adults (Elble, 2003, 2004; Mergl, Tigges, Schroter, Moller, &

Hegerl, 1999; Raethjen, Pawlas, Lindemann, Wenzelburger & Deuschl, 2000) as well as adults experiencing physiological, cognitive, or emotional stress and anxiety (Hung et al., 2005; Wing & Baddeley, 1979).

Wiggle has also been observed in clinical populations, including individuals with Parkinson's disease (Van Gemmert, Teulings, Stelmach, 1998), Alzheimer's disease (Burton, Strauss, Hultsch, Moll, & Hunter, 2006), multiple sclerosis (Longstaff & Heath, 2003), schizophrenia (Caligiuri, Teulings, Dean, Niculescu, & Lohr, 2009), drug-induced parkinsonism (Caligiuri, Teulings, Filoteo, Song, & Lohr, 2006),and essential tremor (Ulmanová et al., 2007). Wiggle has been explained as an adaptation to stress or disease characterized by increased biomechanical stiffness during the writing process, which results in decreased smoothness and increased short stiff strokes in curved trajectories (Longstaff & Heath, 2003; Van Den Heuvel, Van Galen, Teulings, Van Gemmert, 1998; Van Galen and Schomaker, 1992; and Van Gemmert and Van Galen, 1997).

Researchers claim that wiggles reveal important information about the state of the writer (Djebbari et al., 2004; Evidente, 2000; Louis, Wendt, and Ford, 2000; Longstaff & Heath, 2000; Qu, 2004; Vincent, Bouletreau, Sabourin, and Emptoz, 2000). Wiggles may provide information about the underlying dynamics of perceptual-motor coupling during writing (Carson, 1996; Kandel, Orliaguet, & Boë, 2000; Sita, Rogers, & Found, 2003), and features of the writing task such as the type of instrument used for writing, the paper, the writing surface and variations in speed and pressure (Vincent et al., 2000; Wann and Nimmo-Smith, 1991).

Plamondon and Srihari (2000) reported that handwriting fluency depends upon highly skilled anticipatory control of the writing trajectory. Cautious anticipatory control results in slower handwriting speeds and increased wrinkliness in forged signatures and false written statements (Chen et al., 2003; Dick, Found, & Rogers, 2000; and Van Gemmert and Van Galen, 1996). Very slow writing speeds can also contribute to a "wiggle" in the writing line. Sanguineti, Frisone, Bruni, & Morasso (1998) claimed that unnaturally slow speeds produced loops that appeared to be a concatenated series of stiff subsegments rather than a fluid set of curved lines, which appear as wiggles in the writing line. Thus, Vincent et al. (2000) discourages any smoothing of the writing curve, reporting that such methods "make the image independent from the writer and … induce an important loss of information" (p. 87).

Few experimental studies manipulating variables to assess the effects on smoothness in handwriting or drawing movements have been published. In one such study where the subjects' attention was directed to the output of their pen motions, more disturbances in the writing were observed (Marquardt, Gentz, & Mai, 1996).

2.8 Data Selection

2.8.1. Selecting a Segment of the Written Trajectory for Analysis

Methods for analyzing segments of the recorded traces of continuous writing and drawing movements are as varied as the types of handwriting research conducted (Plamondon & Srihari, 2000). Choosing segments for analysis, termed graphical units, is a challenging task because it can be very difficult to determine precise boundaries between one graphical unit and another (Plamondon & Srihari, 2000).

The choice of method for segmenting the recorded trace may be based on theoretical considerations. Studies involving information processing paradigms, such as research on handwriting identification and computer recognition, motor learning, and cognition often use complete letters, syllables, or words as the unit of analysis (Schomaker, 2008; Teulings, Thomassen and Van Galen, 1983). Scientists exploring the dynamics of handwriting, including dynamicists, computer scientists, neuroscientists, and some cognitive psychologists are more likely to divide the recorded trace into smaller curves, based on Hollerbach's 1981 oscillation theory, kinematic theory (Plamondon, Alimi, Yergeau, & Leclerc, 1993; and Woch & Plamondon, 2004), 2/3 power law (Dounskaia, 2007; Plamondon & Alimi, 1997; Richardson & Flash, 2002; Viviani, 2002), underlying force dynamics (Thelen, Corbetta, & Spencer, 1996) or the psychomotor model of Van Galen (1991). For example, Brault and Plamondon (1997) based their curve segmentation procedures on the perceptual importance of points on a curve. Wing (1978) defined pairs of upstrokes as one graphical unit, while other researchers have used spatial or temporal factors, acceleration profiles, maximum and minimum curvature, and minimum and maximum absolute velocity to select endpoints for graphical units (Maarse and Thomassen, 1983).

However, theoretical factors are not the only consideration when selecting endpoints within curved trajectories to divide the sample into graphical units for analysis. Since most segmentation is performed by computer, subroutines must be programmed that enable the computer to recognize the relevant curves, divide them in a consistent manner, and match the appropriate points for comparison. Consequently, some methods for segmenting written curves for analysis arise from practical or mathematical considerations necessary for developing a successful segmenting computer program, rather than theoretical factors alone. A sampling of methods chosen for sampling includes these identifying features: (a) the minima in a curved trajectory (Bulacu & Schomaker, 2007); (b) points of maximum curvature (Morasso, 1986); (c) absolute velocity minima (Schomaker & Teulings, 1992); (d) zero vertical velocity points (Karls, Maderlechner, Pflug, Baumann, Weigel, & Dengel, 1993; Singer & Tishby, 1994), and (e) vertical *y*(t) maxima or *y*(t) minima (Hayes, 1980; Hollerbach, 1981; Liao & Jacaginski, 2000).

Of particular interest in this study, Liao and Jagacinski (2000) chose peaks and troughs of curves (specifically *y*-maxima and *y*-minima) for segmentation. This choice may be especially appropriate when investigating stability, since Roerdink, Ophoff, Peper, and Beek (2008) have found that hand motions extending in a vertical direction are most consistent at reversal points found at *y*-maxima and *y*-minima. These points, which they named anchor points, appear to stabilize the rest of the cycle. Since spatial and temporal consistency was investigated in the present study, I adapted the method applied by Hollerbach (1981) in which loops were divided at the *y*(t)-maxima reversal points at the peaks of the major axes of the loops. Thus the curve between the two peaks was used for spatial and timing comparisons to take advantage of the stability at those points.

2.9 Experimental Apparatus and Software

In this study, I used a Wacom digitizing tablet with an electronically engineered pen which allowed for concurrent real-time feedback of the movement trajectory on a computer display (Plamondon & Srihari, 2000). Such equipment provides "two-dimensional coordinates [or three dimensional coordinates, if pressure information is collected] of successive points of the writing as a function of time…stored in order producing a spatio-temporal representation of the input" (Plamondon & Srihari, 2000, p. 64). I collected spatio-temporal signals from participant writings with the Optimized Action Sequence Interpreter System (OASIS) software developed by Kikosoft, Inc.

(Doetinchem, The Netherlands) and operated with the OASIS Box, a custom-designed machine for running OASIS in DOS.

Researchers using the OASIS software system have studied a wide range of psychomotor disorders through hand-directed pen movements, including Parkinson's Disease (Van Gemmert, Adler, and Stelmach, 2003), Major Depressive Disorder (Sabbe, Hulstijn, Pier, & Zitman, 2000), Multiple Sclerosis (Longstaff and Heath, 2006), and Developmental Coordination Disorder (Kagerer, Contreras-Vidal, Bo, & Clark, 2006; Smits-Engelsman, Wilson, Westenberg, & Duysens, 2003). *OASIS* has been used to assess the kinematics in deception and forgery (Black, Found, & Rogers, 2003; Van Gemmert, 1999) and in psychopharmacological tests (Morrens, Wezenberg, Verkes, Hulstijn, Ruigt, Sabbe, 2007; Sennef, van Riezen, de Jong, & Hulstijn, n.d.). Thus, this technology has been subjected to rigorous testing with a wide range of conditions and populations, and considered to be a reliable instrument for collecting data.

2.10 Data Analysis

In this study, I sought to explore the effect of concurrent visual feedback on timing accuracy, spatial accuracy, and smoothness of the writing line while tracking a target with constant and changing speeds. Individual participants who were presumed to be representative of the larger population of college students (i.e., not of intrinsic interest) were measured repeatedly during each trial and in different trial conditions.

The data generated from this study was correlated, unbalanced, nonlinear, and demonstrated non-constant variance, which produced a challenge for effective and precise statistical analysis.

In many cases, multiple measures taken on the same units or individuals can be analyzed with repeated measures designs. However, repeated measures ANOVA is not the preferred approach when data are unbalanced, heteroscedastic, or both, or when there are few participants measured in relation to the total number of observations available for analysis (Lindstrom & Bates, 1990). Further, studies have shown that repeated measures ANOVA introduce bias, lack power, and inflate standard errors and p-values when data are unbalanced or heteroscedastic (Zuur, Ieno, Walker, Saveliev, & Smith, 2009).

Mixed models, also termed random effect models, random coefficient models, hierarchical models, and multilevel models, were developed from regression models to assess data that failed to meet the assumptions for traditional ANOVA, particularly when data are correlated or hierarchically arranged (i.e., nested) in repeated measures, longitudinal, and blocked designs, and when the specific individuals measured are not of intrinsic interest. Mixed effect models are structured to account for fixed effects from categorical variables of interest and random effects from multiple measurements (Zuur, Ieno, & Smith, 2007). They are "a relatively recent statistical development (e.g. Wang, 1998; Verbyla et al., 1999; Lin and Zhang, 1999), [which] allow for smooth functional relationships, subject-specific effects, and time series error structure" (Coull, Schwartz, & Wand, 2001, p. 339). Since mixed models contain both fixed and random variables, they preserve power and information that are lost in repeated measures ANOVA (Baayen, Davidson, & Bates, 2008) and reduce bias (de Leeuw & Meijer, 2007) that is introduced when measurements are taken multiple times on a subject or unit of interest.

However, linear mixed effects models do not always resolve all the problems encountered by their traditional ANOVA predecessors when data are nonlinear or have non-constant variance. Consequently, statistical researchers have more recently developed generalized additive mixed effects models (GAMM) for analyzing nonlinearity in data that either cannot be sufficiently resolved by transformations and interactions or should not be transformed because the nonlinearity is of scientific importance (Zuur et al., 2007; Zuur et al., 2009). Generalized additive mixed model (GAMM) structures extend mixed model technology by applying smoothing methods to enable effective analysis of nonlinear features of the data. Thus, mixed effect structures including LME and GAMM models are particularly useful in investigations of within-subject differences when data are not normally distributed, or when they are correlated, nonlinear, have non-constant variance, or have missing observations. In short, these models are well-suited to messy behavioral data.

Given these advantages, I used linear mixed effect (LME) models to analyze wiggle data in each sample, and generalized additive mixed models (GAMM) to assess timing and spatial differences between the target and the participant. Each of these models will be described in greater detail in Chapter 4.

2.11 Summary

In this chapter, I have outlined the literature that supports my premise that handwriting research is an area that contributes to our understanding of cognitive psychology and motor control and has continued relevance in a digital age. I began by reviewing the history of handwriting research and various medical, psychiatric, and pharmacological investigations into handwriting. Then I reviewed findings from recent handwriting studies that have developed with advances in computing and research in cognitive sciences and dynamical systems. I followed this section with a review of the debate over theories of handwriting, focusing upon features of information systems and dynamical systems theories, such as motor programs, coordination, attractor states, and synchrony, and discussed controversies over the role of concurrent visual feedback during handwriting tasks. Finally, I discussed the literature that contributed to the experimental tasks, equipment, and statistical methods chosen for this study.

In Chapter 3, Methods, I will provide further details to explain how the study was conducted, followed in Chapter 4, Statistical Analysis, by a complete discussion of the statistical methods used to analyze the data obtained in this study.

CHAPTER III: METHODS

Chapter 3, *Methods*, contains seven sections, beginning with a description of the experimental task and the research design in the first section. A discussion of the characteristics and selection of Participants follows in the second section. The third section of this chapter contains details of the experimental apparatus used with Participants. The target stimuli and testing procedures are discussed in sections four and five, respectively. The predictor (independent) and outcome (dependent, or response) variables are outlined in section six. The chapter concludes with an overview of the analyses conducted on the data in the seventh section.

3.1. Goals of the Present Study

In this study, I sought to explore the effect of concurrent visual feedback on characteristics of timing accuracy, spatial accuracy, and smoothness of the recorded traces drawn by participants as they attempted to couple their hand motions with the movements of a loop-drawing visual stimulus appearing on a computer monitor. The effect of concurrent visual feedback in these sets of tasks is an empirical question that forms that basis of this study.

3.2. Experiment

3.2.1. Experimental Task

In a set of experiments that varied both the feedback and the speed of tracking tasks, 35 healthy right-handed participants used a stylus on a digitizing tablet to track a left-to-right loop-drawing animation presented on a computer monitor. A dot target moved over a template of twelve connected cursive letter *e's*, leaving a track as it drew over each loop in the series. Participants were instructed to draw along with the target to reproduce the shape of the loops at the tempo of the target. Specifically, participants were instructed to keep the pen as close as possible to the leading yellow dot of the target stimulus, matching the tempo and trajectory of the stimulus as closely as possible. Participants performed sixteen trials in a 2 × 2 design, eight trials with their trace visible on the computer monitor and eight trials without a visible trace, half with a constant target rate and half with a variable rate change mid-trial.

3.2.1.1. Experimental Conditions

Four conditions were randomly presented in a 2x2 design to participants:

(1) The recorded trace was not presented on the screen concurrently with the target, which moved across the template at a constant rate, referred to as the Feedback:Absent-TargetRate:Constant condition.

(2) The recorded trace was not presented on the screen concurrently with the target, which moved across the template at a changing rate, referred to as the Feedback:Absent-TargetRate:Changing condition.

(3) The recorded trace was presented on the screen concurrently with the target, which moved across the template at a constant rate, , referred to as the Feedback:Present-TargetRate:Constant condition.

(4) The recorded trace was presented on the screen concurrently with the target, which moved across the template at a changing rate, referred to as the Feedback:Present-TargetRate:Changing condition.

The changing rate was introduced in half of the trials in order to assess whether the effect of the concurrent visual feedback would be robust in the presence of a timing perturbation.

3.2.1.2. Outcome Measures

Three outcome measures were explored: (1) spatial matching errors and (2) timing errors which reflected tracking accuracy and (3) smoothness (wiggle).

Spatial errors. My goal was to observe and note copying errors between the target shape and the shape drawn by the participants. The spatial errors were relative to shape of the target, but not position, since it would be expected that participants would misalign their drawings in the Feedback:Absent conditions. Spatial errors were determined by calculating the root mean square error, RMSE, which will be discussed later in this chapter.

Timing errors. Timing errors were observed for the differences between the times that the target and the participants reached corresponding landmarks along the sample. In addition, I observed how constant and irregular target movement affected these errors to determine whether the effect of concurrent visual coupling was robust to timing perturbations. Thus, I was able to investigate the effect of a visual stimulus on the consistency and coupling of timing between the participant pen-movements and the oscillating target.

Smoothness errors or "wiggle". The third outcome measure of this study is decreased smoothness (i.e., slight undulations) in the writing line, represented by short, stiff, rapidly alternating strokes, forming the appearance of a "wiggle" in the recorded trace. The percentage of wiggle produced in each sample under each of the four experimental conditions was measured to determine how the amount of wiggle varied in the presence or absence of concurrent visual feedback or with constant or changing target rates.

3.2.1.3. Research Design

The experiment was conducted as a generalized randomized block design with repeated measures, (also termed a within-subjects design) with individual participants as blocks, using a two-way factorial crossed treatment structure (Feedback with two levels [Absent and Present] and Target Rate with two levels [Changing and Constant]). Each participant was presented with sixteen loop-drawing tasks consisting of each of the four experimental conditions (i.e., FeedbackAbsent:ChangingTargetRate; FeedbackAbsent:ConstantTargetRate; FeedbackPresent:ChangingTargetRate; and FeedbackPresent:ConstantTargetRate). These experimental conditions were presented four times in random order, except for two initial orientation trials that consisted of two trials performed with FeedbackAbsent:ConstantTargetRate. Most participants performed four trials under each of the four experimental conditions; several trials were lost due to researcher error or equipment malfunction.

3.2.1.4. Participants

Seventy-two undergraduate students at the University of Minnesota took part in this study. Students in kinesiology courses were recruited according to a script approved by the University of Minnesota Institutional Review Board. Thirty-one volunteers (2 left-handed females, 2 left-handed males, 23 right-handed females, and 4 right-handed males) participated in pilot studies to test various aspects of the experimental procedure. An additional forty-one volunteers participated in the final study, including 23 right-handed females, 12 right-handed males, 4 left-handed females, and 2 left-handed males. Since the sample of left-handed participants was insufficient for the purposes of making reliable comparisons, data from the six left-handed participants were not included. Thus, data from only 35 right-handed participants were included in the final analysis. Participants were identified by age, gender, and handedness.

Inclusion/Exclusion Criteria. Participants met criteria for participation in the study if they were able to hold a pen comfortably in a dynamic tripod grip, had normal or corrected vision that allowed them to see the computer monitor clearly, and were able to identify target shapes and colors. Two participants reported some degree of color blindness, but stated that they were able to differentiate the colors on the screen, even though the perceived colors were not precisely the same hues as seen by other participants. Three volunteers were excluded from participation because they were not able to perform the research tasks: one volunteer wore a cast on her arm that interfered with free writing movement, and two volunteers were excluded due to an equipment malfunction.

Research Compensation. When volunteers came to the lab to take part in the study, they were informed that extra credit would be awarded to them for coming to the lab, whether or not they chose to participate in the study. The offer of extra credit was not attached to any requirement to perform the experimental tasks so that participants could discontinue their involvement in the study at any point without penalty. The three excluded volunteers received extra credit, despite not being able to participate. To ensure that the participants understood the voluntary nature of their participation, extra-credit verification was emailed to their instructors as soon as they arrived at the lab, before they observed a demonstration of the equipment and the research tasks. The researcher offered volunteers the opportunity to ask questions and to obtain answers before inviting them to participate. Participants completed consent forms developed for this experiment and approved by the University of Minnesota Institutional Review Board.

3.3. Materials and Apparatus

3.3.1. Equipment and Software

This section provides a description of the equipment used in this study, including the digitizer, software and the target stimuli that participants were asked to reproduce.

3.3.1.1. Digitizing Tablet and Pen

The current study utilized a Wacom UltraPad digitizing tablet measuring 77.7 cm (width) by 61.5 cm (depth) and a specially engineered pen (Wacom Technology Company, Wacom Europe GmbH, Krefeld, Germany), as shown in Figure 2. The digitizing tablet provides the surface on which participants reproduce figures shown on the monitor. An electrostatic feature of the tablet prevents paper from shifting during the experiment. The tablet can reliably detect pen movements as small as 0.01 mm or 1000 lines per cm (Ketcham, Dounskaia, & Stelmach, 2006) within the range of speeds performed on the tablet during this experiment, and sampled data at 170-207 Hz during each trial.

According to the technical information provided on the Wacom Technology Company web page (Wacom Technology Company, 2007), the digitizing pad and specially engineered pen operate with Electro-Magnetic Resonance (EMR®), a position-sensing technology, patented by Wacom Technology Company. EMR® couples electric and magnetic fields of the pen and pad, allowing movements of the pen to be detected by circuitry in the digitizing tablet. The pen is free of restrictive electric cords or batteries that can distort normal writing movements (Figure 2).

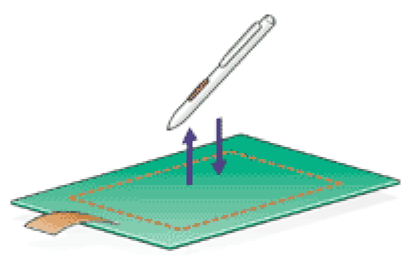

Figure 2. Coupled magnetic and electric fields between digitizing tablet and electronic pen. Copyright 2007 by Wacom Technology Company. Reprinted with permission.

Wacom Technology Company provides this description of the operation of the electronic pen and digitizing tablet:

> Weak energy is induced in the pen's resonant circuit by a magnetic field generated by the sensor board surface. The pen's resonant circuit then makes use of this energy to return a magnetic signal to the sensor board surface. By repeating this movement, the board detects information on the pen's coordinate position and angle… speed and writing pressure…. A sensor unit is equipped at the side of the sensor board to switch the magnetic field on and off and to receive signals at high speed. (Wacom Technology Company, 2007).

Electric and magnetic fields are coupled when the pen is held within 14 mm above the surface of the digitizing tablet, which allows movements of the pen to be detected when the pen is lifted slightly from the tablet. This feature is helpful when monitoring positioning movements of the hand where the pen is not in contact with the paper. Qu (2004) describes the coupling function as follows:

> the coupled energy resonates with the tank circuit [embedded in the pen, as shown in Figure 2] and reflects back towards the sensor board by forming a shaped h-domain field at the tip of the pen. As this happens, [an] antenna coil is switched to receive this reflected energy and provide an analogue signal. This process is repeated in rapid succession with all antenna coils. All of these analogue data are then collected and converted into digital signals that can be

post-processed to give *x* and *y* position and pen tilt [and pressure] information (pp. 28-30).

The pen movement-sensor circuitry in the electronic pen was modified further by Peter de Jong of Kikosoft, Inc. to register pressures up to 800 grams, which are typical in some clinical populations. The circuitry embedded in the pen is shown in Figure 3a. The assembled pen, resembling the heft and size of a standard pen with a moderate-sized barrel, is shown in Figure 3.b.

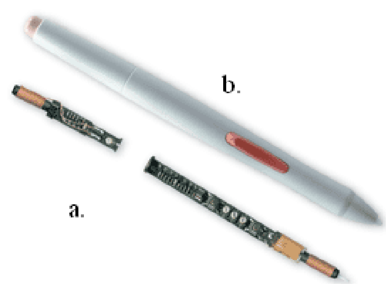

Figure 3. Wacom electronic pen used with digitizing tablet. (a) Interior circuitry, (b) external shell. Copyright 2007 by Wacom Technology Company. Reprinted with permission.

3.3.1.2. Software: OASIS

Spatio-temporal signals from participant writings were collected with the Optimized Action Sequence Interpreter System (OASIS) software developed by Kikosoft, Inc. (Nijmegen, The Netherlands). The data collection and control program was written using the programming language *OASIS* (de Jong, Hulstijn, Kosterman, & Smits-Engelsman, 1996).

OASIS system operation. The OASIS system collects signals from the digitizing tablet, allowing handwriting motions to be sampled from the digitizing tablet at 170-206 Hertz, depending on the computer, the type of tablet, the driver, and the number of variables computed by the OASIS software. The driver component in OASIS allows higher accuracy and higher sampling frequencies than the standard driver provided by the manufacturer of the digitizing tablet. Signals were filtered with a Butterworth filter installed in the software and set to low pass at 8 hertz to remove noise that would result in speed traces that were not smooth, since noise has been found in many studies of tracking behavior (Hick, 1948; Miall, Weir & Stein, 1988; Noble, Fitts, & Warren, 1955; Roitman, Johnson & Ebner, 2001, Searle and Taylor, 1948; and Vince 1948).

The data collected in the current study included *x*-coordinates (PenX) and *y*-coordinates (PenY) of the pen position on the digitizing tablet, a composite pen speed variable that measures speed in both *x*- and *y*- directions(PenV), axial pen pressure (PenZ) exerted on the digitizing tablet, and exact times at which data were recorded

(PenT). Since these quantities were sampled at 170-206 hertz over the course of each trial, movement data were recorded every 0.0049-0.0059 seconds. In most trials, the *OASIS* system sampled pen movements at rates ranging from 175-185 hertz, or approximately every 0.0055 seconds. Faster speeds were recorded late in the study when unnecessary variables were removed from the data collection function.

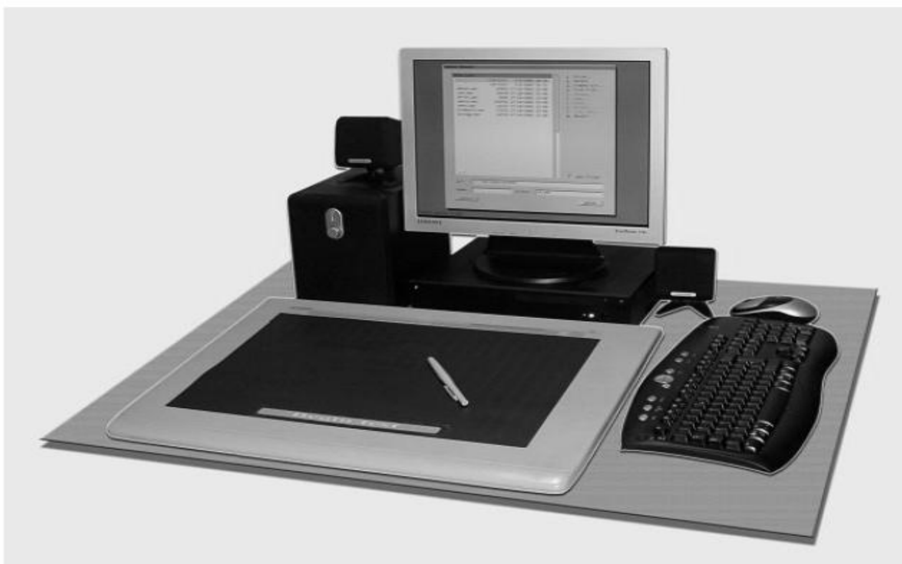

Figure 4. OASIS equipment with Wacom Ultrapad digitizing tablet and electronic pen. Copyright 2007 by Kikosoft, Inc. Reprinted with permission.

3.3.2. Target Stimuli

3.3.2.1. Loop design

Although OASIS comes packaged with standard target stimuli, most researchers program their own stimuli to meet the requirements of their research. The target stimulus loops used in this study were formed by developing algorithms in *Mathematica*® (Version 5.2) to create Bezier curves that resembled the cursive letter *e* by Stan Wagon, Professor of Mathematics and Computer Science at Macalester College, August, 2005), as shown in Figure 5. The Bezier curves configured for this study were replicated twelve times and joined to form a connected sequence, as shown in Figures 6 and 7.

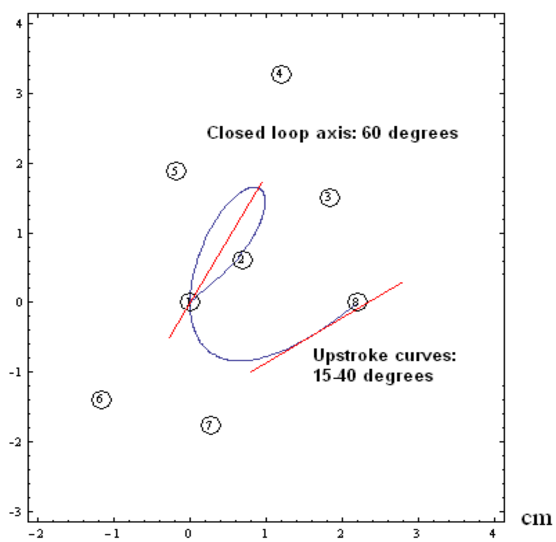

Figure 5. Prototype of the ellipse (blue) replicated to form target stimuli. Circled numbers refer to the eight control points required to form the Bezier curve. Red lines represent the major axis at 60° and the tangent to the connecting upstroke at 25°.

The resulting target was a series of twelve joined *e*'s (***eeeeeeeeeeee***) aligned with the major axis of the loop portion oriented at 60°. This orientation was intended to help writers avoid overlapping adjacent enclosed portions of the *e*'s in the sample, and was recommended by Sylvie Athènes, researcher of human factors and handwriting at Université Paul Sabatier, Toulouse, France (personal communication, May 19, 2005). The tangents on the longest segment of the target, the upstroke, were oriented from 15°-45° along the curve from the trough between adjacent loops to the following enclosed loop. Each loop with initial and terminal strokes measured 2.21 cm (0.87 inches) in the horizontal direction and 2.48 cm (0.98 inches) in the vertical direction.

Following Sallagoïty, Athènes, Zanone, & Albaret (2004), the enclosed portion of the loop was modeled on the shape of the ellipse found to have properties of an attractor. In the present study, the ellipse measured 1.76 cm (0.69 inches) along the major axis and 0.6 cm (0.24 inches) at the minor axis at the widest point of the enclosed loop, approximating the ellipse Sallagoïty et al. (2004) depicted that had a 6:2 ratio between the major and minor axes, respectively. A complete target of twelve connected *e*'s extended 26.5 cm (10.44 inches) in the horizontal direction, and 2.48 cm (0.98 inches) in the vertical direction across the template.

As shown in Figure 5, the major axis of the enclosed loop of the target is oriented to 60°, and the tangents to the upstroke to the following enclosed loop range from 15°-

40° arising from the trough and extending to the widest part of the following loop. The equations forming the curve from the Bezier algorithms were written into the *OASIS* software as macros to form two experimental stimuli:

(1) A static template of twelve connected loops consisting of blue dots was used as a pattern for tracing and as an orientating mechanism for the loop-drawing task, as shown in Figure 6.

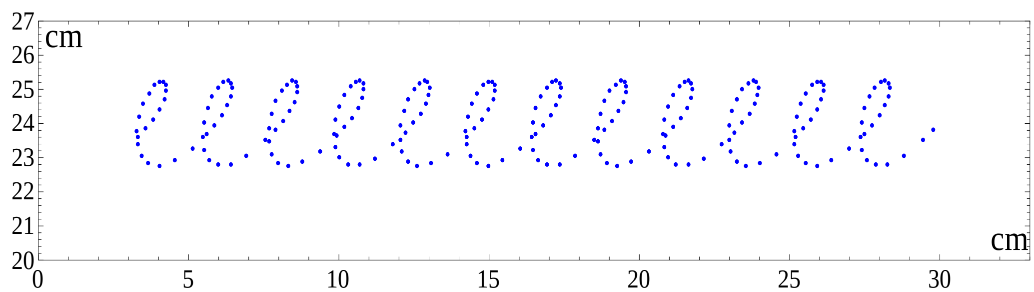

Figure 6. Static template of blue dotted loops built from the prototype for the shape of the *e*'s shown in Figure 5.

(2) A moving yellow dot traced over the blue dotted template, covering the blue dots on the template with a solid red line (Figure 7).

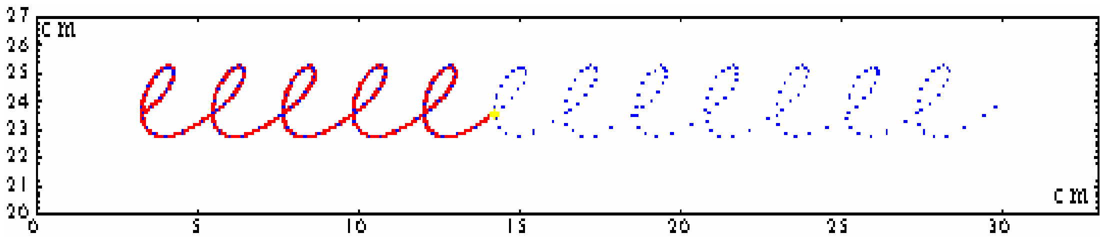

Figure 7. Yellow dot tracing over static template of blue dotted loops, covering the blue dots with a solid red line as the dot traverses the template.

3.3.2.2. Target Rate

In a pilot study preparatory to the current research, participants were asked to draw a series of twelve loops in several different formats to determine the optimum target speed, including these:

(1) A series of twelve loops were drawn freely at the participants' own pace without a template or animation to guide their movements.

(2) A series of twelve loops were drawn over a static template without a moving target.

(3) A series of twelve loops were drawn by an animated target without an underlying template.

(4) A series of twelve loops were drawn in sequence by an animated target overwriting a static template consisting of twelve loops.

Participants expressed a strong preference for the fourth format (i.e., 4) as the one they found most conducive to frequency and spatial coupling. They commented that they were not able to maintain a consistent performance at a natural tempo without seeing both the static template of twelve loops, which appeared on the screen before the animated target appeared, and the animated target moving over the template. Without the animation, the participants drew at unnaturally slow rates to ensure careful tracing of the template loops in the conditions where feedback was present. Without the static template, participants remarked that they did not know how to plan their movements since they could not predict the trajectory of the target. The participants reported that their drawing movements felt most natural when the static template was shown as it was overwritten by a dynamic target. The presence or absence of concurrent visual feedback made no difference in their judgments.

The reflections of those who participated in the pilot phase of the present study are consistent with several studies. For example, Kaiser et al. (1992) reported that animations improve the ability of the observer to predict and track target trajectories over the use of static diagrams. A study by Orliaguet, Kandel, and Boë (1997) found similar results, leading these researchers to conclude that moving experimental stimuli enabled more accurate predictions of subsequent target trajectories. However, the animated target alone was not sufficient for the pilot participants in this study. Participants commented that they needed the static pattern to determine the planned trajectory of the animation, even though they benefited from the dynamic information from the target. Participants commented that the static template allowed them to estimate the size and spatial alignment of the target, while the animated target enabled them to predict the dynamics of the moving target, allowing them to initiate movements with less delay and maintain frequency coupling as closely as possible.

Given the results from these two studies and participant comments, the target stimulus was created with two components:

(1) A static, dotted pattern of twelve loops was presented before and during the trial.

(2) An animated target consisting of a yellow dot covered the loops on the template with a solid red line as it traced over them, as shown in Figure 7.

The target stimulus in the present study appears to be novel for handwriting and loop drawing studies, since previous literature has not used a combination of static and animated patterns as a visual stimulus in a pursuit loop-drawing experimental task.

3.4. Segments Selected for Comparison and Analysis

The method selecting segments for analysis in this study was an adaptation of an approach used by Hollerbach (1981) to determine writing slant. Hollerbach's method consisted of drawing a line horizontally from the first peak (or cusp) of the letter *u,* located at a *y(t)* maximum, to the following peak, as shown in Figure 8.

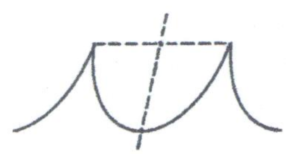

Figure 8. Method of establishing the orientation of the primary vertical and horizontal axes of *u*-shaped letters described in Hollerbach, J.M. (1981), p. 142. Reprinted with permission.

Using the method described by Hollerbach (1981), the *y(t)* maxima coordinates for the target and associated participant writing were divided into peak-to-peak segments to form the basic units for comparison. A peak-to-peak interval is shown as the curve spanning the trajectory between two successive peaks marked by green dots in Figures 9 and 10. In a sample template of 12 loops, 11 peak-to-peak segments span the sample. This method was used to select portions of the curves for analysis in the present study because of precision in identifying consistent end points for each measured loop. In trials where participants drew over previously drawn loops or navigated off the template, as occurred in trials where the participants were not provided with a visual trace of their drawings (Figure 10, for example), identifying endpoints with any other method was particularly unreliable.

In the present study, peaks are defined as *y(t)* maxima in a loop; they are shown as green markers in Figures 9 and 10.

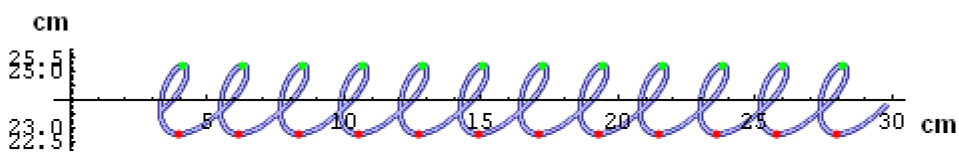

Figure 9. Sample of twelve loops (blue) drawn over the static template (not visible). Loop peaks (i.e., maxima of *y(t)* coordinates) are indicated by green markers; loop troughs (i.e., minima of *y(t)* coordinates) are indicated by red markers. This example shows a perfect spatial matching to the target stimulus. The participant line is shown in blue in this example, covering the red line left by the target as it traversed the template.

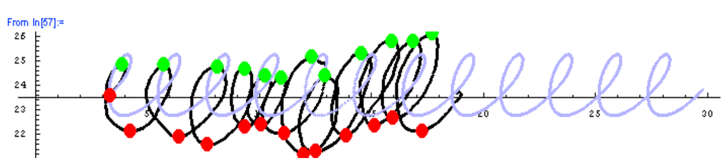

Figure 10. Participant sample produced without concurrent visual feedback (i.e., Feedback:Absent conditions). Target is represented by light blue loops; recorded trace by participant (black line) contains green markers at loop peaks and red markers at loop troughs. Markers are increased in size in Figure 10 for improved readability.

Using segments of the trajectory that spanned the length of the curve from the point(s) at the peak-to-subsequent-peak emerged as the best solution to several segmentation problems, enabling a reliable and precise method to identify endpoints to match curves between the target and participant drawing. Using troughs (y(t)-minima) was not acceptable because the lowest points on the curve forming the loops were not consistently found at the midpoint between adjoining loops. Even when the lowest points in the trajectory occurred in the trough between loops, the trajectory sometimes was horizontal for a short range in those locations, producing a large number of points with the same value on the y-axis over a span of points on the x-axis, requiring an arbitrary decision regarding the location of the segment endpoint. As a result, the shapes and lengths of loops extending from a $y(t)$ minimum to a subsequent $y(t)$ minimum varied considerably and did not produce segments that could easily be compared. Thus, peak-to-peak segmentation at $y(t)$ maxima yielded results that were far superior to other curve matching methods attempted in this study.

This method of segmenting by selecting peak-to-peak segments is less common than segmentation at zero velocity crossings (Plamondon and Guerfali, 1996; Plamondon and Alimi, 1997). Using the zero-crossing approach made it difficult to compare similarly shaped curves, as changes in acceleration and velocity were not consistent across conditions. Comparing segments using times as landmarks is another well-regarded approach for segmenting curves for comparison which matches locations on the curve at which the participant and target were found at the same times also produced differently shaped curves for comparison. Comparing the pen point positions when the target reached specific points in the cycle (peaks or troughs) was equally unsuccessful, because it was similar to comparing unsynchronized clocks running at different rates. Any coupling at peaks that occurred between the participant pen point position at a certain time in the cycle and the position of the target seemed to occur by chance alone. Once a match between the participant point in the cycle and the target occurred, coordination was not maintained and the timing relationships between the participant and the animation quickly diverged, so that the participants' motions did not remain synchronized with the target.

Even though the method of using *y(t)* maxima is less common than the zero-crossing approach, segmenting with *y(t)* maxima and *y(t)* minima has been used in other studies. For example, a similar method of segmenting was used by Liao and Jagacinski (2000) in a manual tracking task. Ketcham, Dounskaia, and Stelmach (2003) provided theoretical support for selecting peaks for segmenting written curves for analysis by noting that

[i]n recent work by Reina and Schwartz (2003) on a cyclical oval-drawing task in monkeys, visual gaze fixated cyclically on the points of highest curvature. Those points of highest curvature may serve as anchor points or, more specifically, may provide the system with necessary information regarding the size and shape of the trajectory. Points

of highest curvature have been of interest for many researchers because those are where the velocity of a movement is the lowest, as stated by the two-thirds power law (de'Sperati & Viviani, 1997; Lacquaniti, Terzuolo, & Viviani, 1983; Viviani & Flash, 1995; Viviani & Terzuolo, 1982) (p. 30).

Roerdink, Ophoff, Peper, & Beek (2008) found that anchors, occurring at reversal points (in this case, *y(t)*- maxima and minima) exhibit lower variability than other points on the drawn cycle, suggesting an organizing role for the points at those locations. Thus, segmentation at *y(t)* maxima appeared to be the most promising option for obtaining consistent results by increasing precision in choosing appropriate end points for selecting comparable curves, and exploiting *y(t)* maxima as stable points for organizing the coordination dynamics of the cyclical drawing.

3.5. Testing Procedures

3.5.1. Experimental Set-Up

In this study, participants were seated at a desk facing a flat 17" computer monitor and an 18" x 24" electrostatic digitizing tablet, as shown in Figure 11. In the figure, the electronic pen is placed on the digitizing tablet, and the monitor shows a menu of *OASIS* functions. Note that the pen is unencumbered by cords, allowing free movement. Stimuli were presented on the monitor, and participants' movements of the pen generated position, timing, and pressure data that were transmitted to the *OASIS* software for analysis.

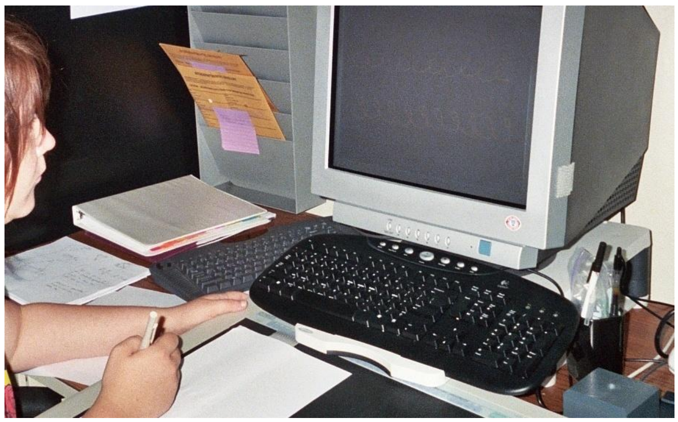

Figure 11. Participant is seated at a desk, watching the target on the monitor, while reproducing the shape and tempo of the target on paper placed on the Wacom® Ultrapad digitizing tablet. She is using the specially engineered non-inking pen to track the target.

3.5.2. Experimental Conditions

Each writing test was divided into two parts: a set of two training trials (Feedback:Absent-TargetRate:Changing) that allowed the participant to become familiar with the equipment, followed by a set of 14 randomized trials consisting of two additional trials in the Feedback:Absent-TargetRate:Changing condition, and four trials each with Feedback:Absent-TargetRate:Constant, Feedback:Present-TargetRate:Changing and Feedback:Present-TargetRate:Constant conditions. Data were collected and analyzed from all 16 trials (the two training trials and 14 randomized trials) performed by each participant.

3.5.2.1. Training Trials

For the training trials, participants were seated comfortably at the desk, facing the monitor. The digitizing tablet was placed between the monitor and the writer with the pen placed vertically in a pen holder that kept the pen erect on the center of the pad. Participants were asked to pick up the pen as they would to begin writing. When the participant lifted the pen from the pen holder, the experimenter removed the pen holder from the center of the pad. The choice of hand used to pick up the pen and the participant’s response to the question “Are you right handed or left handed?” determined handedness for the purposes of this study.

Participants were asked to observe the blue-dotted test template on the monitor in front of them, and to place the pen point at the starting position on the digitizing tablet. This position corresponded to the beginning dot at the far left position on the blue template as shown in Figure 9. A thin arrow was placed on the digitizing tablet to enable the participant to locate the starting point on the digitizing tablet. This point was located approximately 3.5 inches (approximately one handbreadth for the average male) to the left of the participant’s midline. The participant’s chair was adjusted as necessary to ensure proper positioning. Before the trial began, the position of the pen tip was displayed as a thin blue line on the screen, enabling participants to set the pen tip as closely as possible to the precise starting point. Participants were directed to trace over the template while attempting to keep the pen point within the leading yellow dot.

The beginning of each trial was marked by the emergence of a leading yellow dot. As soon as the dot became visible, it proceeded to trace over the blue template, changing the blue dots to a solid red line as it traversed the template. In trials where feedback was provided (i.e., Feedback:Present-TargetRate:Changing, Feedback:Present-TargetRate:Constant conditions), the trace made by the participant appeared as a green line on the monitor. No feedback was provided during the training trials.

Participants were given two practice trials to trace along with the target, moving over the experimental template at a constant rate. In the two training trials, the target moved at a constant rate. In these practice trials, participants did not receive visual feedback from their pen movements while performing the loop-drawing task. The experimenter also asked each participant to name the colors on the screen to assess whether the participants were able to distinguish the moving target from the static template and thus, view the trace generated by their hand movements as they moved to the starting position. At the end of the practice trials, the experimenter said “That's right. Does that feel comfortable to you?” In two cases, participants reported some degree of color blindness, but stated that they were able to distinguish the target and the writing trace generated when they moved into position, even though the perceived colors were not precisely the same hues as seen by other participants. The experimenter concluded by asking, “Do you have any questions?” If there were no issues of discomfort or questions, the experimenter asked, “Would you like to continue?” If the participant agreed to proceed, no additional instructions were given to participants throughout the remaining fourteen trials.

3.5.2.2. Randomization of Experimental Conditions over Trials

Following the two training trials, each participant performed 14 random trials, for a total of 16 trials. Randomization for the order of the trials for each participant was achieved by selecting fourteen cards in sequence from an opaque box, each identifying one combination of experimental conditions. After the initial training trials, a card representing one of the remaining trials to be performed during the experimental session was drawn after each completed trial, including

(a) two additional trials without visual feedback of the recorded trace on the screen with the target moving at a changing rate (Feedback:Absent-TargetRate:Changing condition);

(b) four trials without concurrent visual feedback with the target moving at a constant rate (Feedback:Absent-TargetRate:Constant condition);

(c) four trials with concurrent visual feedback of the recorded trace on the screen along with the target stimulus moving at a changing target rate

(Feedback:Present-TargetRate:Changing condition), and

d) four trials with concurrent visual feedback of the recorded trace on the screen along with the target moving at a constant rate (Feedback:Present-TargetRate:Constant condition).

For each of the randomized trials, the researcher shook the box to shuffle the cards containing a trial condition, and then selected one card at random. Once the participant completed a trial, the researcher set aside the card with indicating the just-completed trial condition, shook the box again, and selected another card. Thus, each participant performed two training trials, followed by 14 additional trials in a random order, completing four trials in each of the four combinations of experimental conditions.

Each trial condition was initiated from the *OASIS* macro menu while the participant was asked to move into the starting position. A blue line tracking the participant's movements was shown on the monitor to assist the participant in placing the pen tip in the starting position. Participants were instructed to keep their pen tip within the leading yellow dot or as close as possible as it moved across the dotted blue template. When the participant responded positively to the researcher's question "Ready?" the trial was initiated. The trial began when the yellow dot emerged at the starting point and began moving over the template loops.

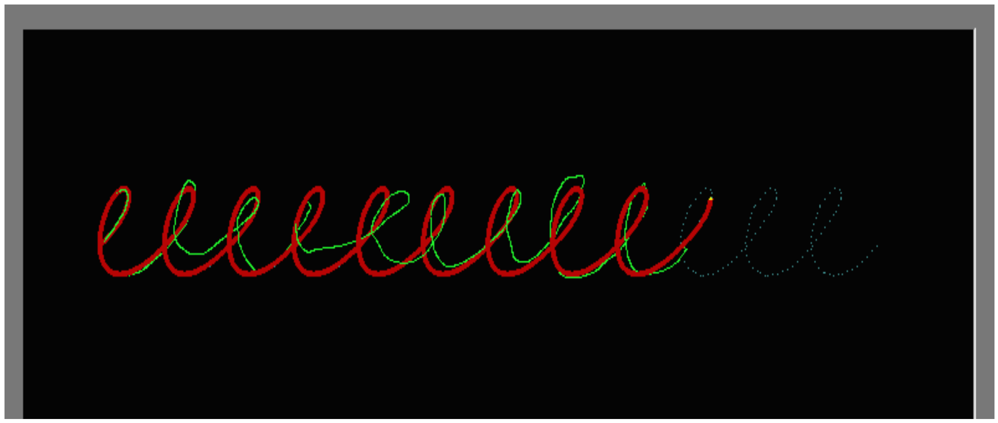

Figure 12. Target moving across blue-dotted template with trace shown as a green line. The underlying blue dotted template can be seen faintly to the right of the animated target stimulus.

3.6. Study Variables

3.6.1. Predictor (Independent) Variables

Differences in timing, spatial arrangement, and irregularities in the written line ("wiggle") that occurred under the (a) the presence or absence of a visual trace of the writer's pen movements, and (b) constant or changing rates of movement of a target stimulus were explored, which were combined into the four experimental conditions discussed previously.

3.6.1.1. Concurrent Visual Feedback

A blue line tracked the participant's pen movements prior to the onset of the trial to assist the participant to position the pen tip as accurately as possible at the starting point. The blue positioning line turned to green at the start of the trial. When the trial began, a leading yellow dot began tracing the loops on the template, turning them to red such that it appeared that the yellow dot was leaving a red trail as it moved across the static template.

In eight of the 16 trials, concurrent visual feedback was provided to participants in the form of a visible trace of their own writing line, appearing in real time as a green line that tracked their pen movements from the moment the trial began. Thus, there was no noticeable delay between participant tracking movement and the appearance of the recorded trace on the screen as the writer followed the yellow dot across the template, (Figure 12), providing concurrent augmented visual feedback. In the other eight trials, the visual trace was not provided; Participants were instructed to keep up with the leading yellow dot on the moving target as closely as possible, even though their own writing trace was not visible as a green line on the monitor to guide their performance. No auditory feedback was provided. The target stimulus remained on the screen after the yellow dot completed the trace to allow participants enough time to copy all 12 loops in the target.

3.6.1.2. Changing and Constant Target Rate

The target moved at a constant rate during eight of the 16 trials, and moved at changing rates in loops 3-7 during the other eight trials. In the constant rate condition, the target completed each loop in 2.35 seconds. This rate was chosen based on the mean values for a preferred or comfortable pace as judged by participants in the pilot phase of the study. Thus, for most pilot participants, 2.35 seconds was considered neither too fast nor too slow for the task.

In the TargetRate:Constant conditions, the yellow dot moved over the blue template in 28.3 seconds, but the time to complete the trial was set at 39 seconds so that Participants might be able to complete the entire template. In the TargetRate: Changing conditions, the target began at a constant rate for the first two peak-to-peak segments (2.35 seconds per peak-to-peak segment), then sped up, slowed down, and resumed the original constant rate. The resulting speed patterns were as follows:

(a) 2.35 seconds each from the first peak to the second peak and from the second peak to the third;

(b) 1.89 seconds from the third to the fourth peak;

(c) 0.95 seconds from the fourth to the fifth peak;

(d) 0.66 seconds from the fifth to the sixth peak;

(e) 1.38 seconds from the sixth to the seventh peak; and

(f) a constant rate of 2.35 seconds from the seventh to the eighth peak, and for the remaining segments of the target template.

3.6.1.3. Gender

Gender was treated as an independent variable to determine whether males (n = 12) or females (n = 23) performed differently on the experimental tasks. This aspect of the investigation is underpowered due to small sample size. Therefore, findings must be viewed cautiously.

3.6.1.4. Pen Speed

Pen speeds were correlated with timing errors in the Feedback:Absent-TargetRate:Constant condition as a check on the functioning of the software. Values for pen speed (*PenV*) were computed in the OASIS software by

$$\sqrt{Pen(VX)^2 + Pen(VY)^2} \qquad \text{(Formula 1)}$$

where *PenVX* is the speed of the pen point in the *x*-direction, and *PenVY* is the speed in the *y*-direction. In the current study, a value for *PenV* was obtained at each sampled point on the recorded traces provided by participants, and an average value was computed to yield a global dependent variable for the entire sample.

3.6.2. Outcome (Dependent) Variables

Data were analyzed for three response variables: (a) spatial displacement errors; (b) timing errors; and (c) wiggle in the contour of the participant writing. The coordinates of the pen positions at each sampling time were acquired and computed by the *OASIS* software from signals generated by the Wacom® tablet. These values were imported into *Mathematica*® and computed with programs designed for these purposes.

3.6.2.1. Spatial Errors as Measured by Root Mean Square Error (RMSE)

Displacement errors in the present study were defined as differences in shape between the peak-to-peak segments drawn by participants and the target as calculated by Root Mean Square Error (RMSE), as shown in Figure 13. In this example, the target is portrayed as the red curve extending from the fifth peak to the sixth peak of a segment of the target, with the participant curve in black superimposed on it.

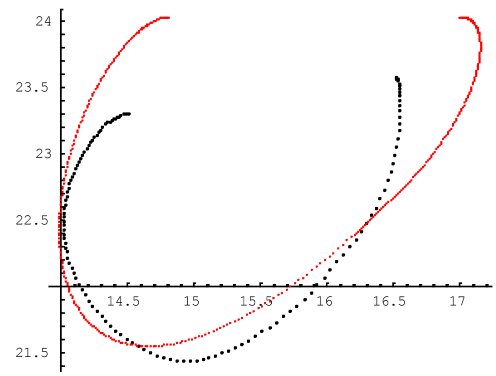

Figure 13. A single peak-to-peak segment from a participant (black) superimposed on the corresponding segment from the target (red).

A measure of the fit between the target and participant's curves was computed using several subroutines coded in Mathematica® to obtain the minimum root-mean-square differences between each peak-to-peak segment of the target and the corresponding segment of a participant's trace. To obtain a root-mean-square-error value for each peak-to-peak segment, each participant point (x_p, y_p) in the segment was fitted to the closest point, measured by Euclidean distances, on the corresponding target curve (x_t, y_t), with no regard for the time difference between the target and participant points. The root-mean-square-error value for the segment was calculated from Euclidean distances between each set of adjacent points on the Participant curve (x_p, y_p) and the closest point on the target curve (x_t, y_t), as shown by Formula 2.

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(x_{p_n} - x_{t_n})^2 + (y_{p_n} - y_{t_n})^2}{n}} \qquad \text{(Formula 2)}$$

After making an initial RMSE error calculation, the program shifted the participant's segment relative to the target curve, and then recalculated the root-mean-square-error value. The process of shifting the participant's curve and computing the root-mean-square difference was iterated until the program obtained the best fit, defined as the minimum root-mean-square-error value (minimum RMSE).

Further checks were added to the program to ensure that (a) upward strokes in the participant sample matched upward strokes in the target sample, and (b) downward participant strokes matched downward target strokes. This additional step was added to reduce matching errors in some samples where the participant loops were drawn much narrower than the target, or where participant loops were drawn on top of each other. Errors were estimated independent of their absolute position to prevent penalizing offset errors in the Feedback:Absent conditions.

Following this method, high spatial accuracy scores were derived as scores with low Root Mean Square Error (RMSE) values and indicated that the recorded traces of the participants closely replicated the shape and size of the target curves, even if their drawing was displaced from the segment being traced. RSMEs for each peak to peak segment in the sample were summed into one error measure for each sample, yielding a summary score that indicated how closely the participants replicated the dimensions of the target shapes.

3.6.2.2. Timing Errors

Timing errors were defined as timing differences between the target and participant drawings at corresponding landmarks (i.e., peaks across the sample). In this study, timing deviations from the target were calculated at specific points, i.e., the peaks of the twelve loops in the sample sequence, defined as the reversal point at *y(t)* maxima for each loop. This approach differs from many studies of timing in coordinated movement, in which temporal consistency is measured within relative phase relationships (see Richardson, Marsh, & Schmidt, 2005; and Volman & Geuze, 2000, for review).

In the present study, temporal coupling was assessed at peaks only, in contrast to relative phase research, which involves relative phase profiles for the entire cycle. Use of peaks for measuring timing relationships provides two advantages in the present study: (a) timing can be measured precisely within 0.006 second and (b) *y(t)* maxima at peaks may locate stable areas on the curve. This view was proposed by Roerdink et al. (2008), who hypothesized that clearly identified reversal points s in the cycle are “anchor points" that function as "intentional attractors" that “make task-specific information . . . available for organizing the cyclical act" (p. 143) in rhythmic tracking tasks capable of organizing timing features of elliptical pen motions.

Timing adjustment procedures. Although each trial began when the computer target began to move over the dotted template, participant timing was adjusted to remove initial jitter and reaction time and the initial differences between the computer generated target motion and participant response by setting the time at which the first peak in each sample occurred to time = 0 seconds. Thus, the first peak in the sequence of target loops and in the participant loops was identified as the starting position. Times in the data set were adjusted to begin the calculation of time differences beginning at the time that the first peak (local y(t)-maxima) of the sample was reached, which was set to zero. Then the differences between the times that the target and the participants reached corresponding peaks were calculated. Consequently, data from the participant times and the target times were adjusted to begin concurrently at the first peak in the sample (Peak 1) to establish a uniform starting position for assessing timing errors.

This approach has been advocated by Ganz, Ehrenstein, and Cavonius (1996) who suggested that removal of initial data removes extraneous features and aids the process of measuring handwriting characteristics more precisely. Thus, this adjustment provides several advantages for calculating timing errors: (a) minimizing extraneous contribution of participant reaction time and (b) reducing start-up effects from the perceptual and effector systems at initial conditions. As a result, each peak-to-peak segment of the participant samples could be compared more accurately to the corresponding segment of the target.

Graphical displays of timing errors. Timing errors were portrayed by a set of plots suggested for this project by Gary Oehlert PhD, Professor of Statistics, University of Minnesota, for displaying differences in timing between the target and the participant trace (personal communication, March 20, 2006). The Oehlert plot approach is an adaptation of a quantile-quantile plot in which one sample is plotted against another to determine whether they arise from the same distribution with similar statistical characteristics (Weisstein, 2008). In this graphical approach, times from the start of each participant trial to each drawn peak were plotted against the times at which the animated target attained the peaks on the template. Although the use of Oehlert plots is a new method for revealing discrepancies in timing between Participant-drawn and computer drawn targets, quantile-quantile plots are widely used to compare data sets (Cook & Weisberg, 1999; Gnanadesikan, 1977).

In Figure 8, the timing errors of one participant in each experimental condition are shown.

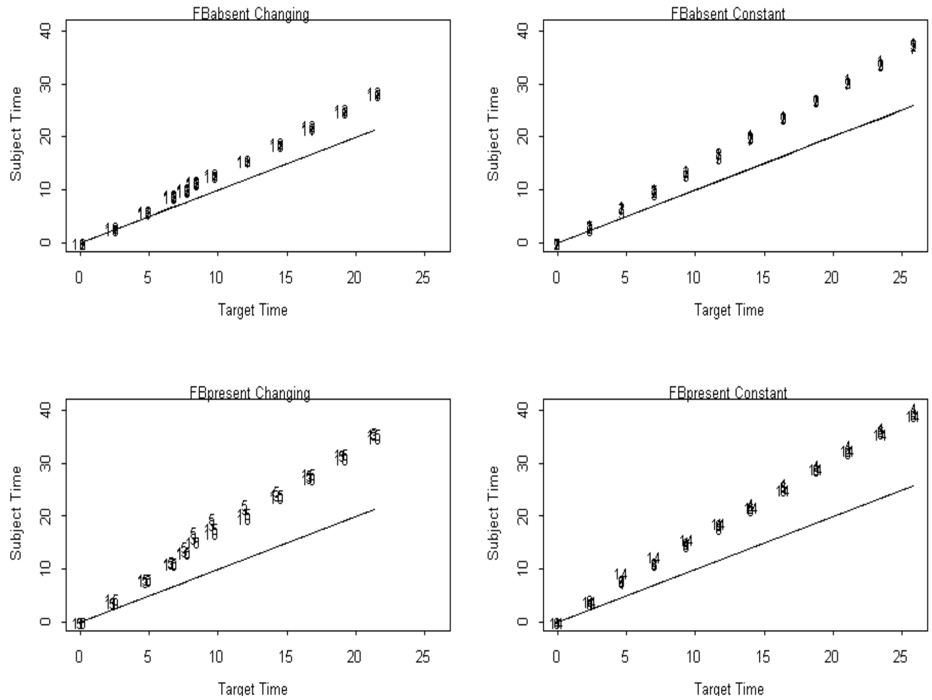

Figure 8. Timing errors of one participant in each experimental condition (Feedback:Absent-TargetRate:Changing, Feedback:Absent-TargetRate:Constant, Feedback:Present-TargetRate:Changing, Feedback:Present-TargetRate:Constant). The numbers indicate the order in which each trial was performed.

The timing errors in Figure 8 show that this participant approximated the errors made at each peak from trial to trial, resulting in trial numbers being superimposed on each other in some areas where errors were very close at the same peaks in different trials. The diagonal line represents the target time across the trials.

As can be seen in the left column of Figure 8, i.e., the TargetRate: Changing conditions, the rate changed after loop 3 through loop 6, and then resumed the original pace. Since the timing in this section of the trial was speeded, trials in the TargetRate: Changing conditions were briefer than in the TargetRate:Constant conditions.

3.6.2.3. Method of Calculating Undulations in the Writing Line (Wiggle)

In this study, the trajectory of a perfectly tracked motion around the loops of the template would tend continuously counterclockwise, and tangents to the trajectory would increase from 0° to 360° around the loop. Disturbances in this trajectory that alternated back and forth over the target line would appear as *wiggles* (also called *wrinkliness* by Chen et al., 2003) and would consist of a series of short strokes with alternating decreasing and increasing angles to the trajectory. For this study, therefore, w*iggle* was defined as the percentage of 21-sampling point segments along the recorded trace where the stroke lifted above the target trajectory, moving in a clockwise direction, over a sample. The resulting correction resulted in an undulation that appeared as a slight “zig-zag” or “wrinkle” in the writing line.

In this approach, tangents were estimated at each point on the writing line, with the exception of the first ten and last ten points of the sample, by secants joining a point located 10 samples previous to the evaluated point to a point 10 samples ahead of the evaluated point. Wiggle was determined at every point on the curve drawn during a trial, except for the ten points at the beginning and the end of the sample. This systematic process allowed very small sections of the drawn curves to be explored for clockwise and counterclockwise motion. Wiggle was detected by locating regions where the curve was decreasing (i.e., tending clockwise) more than 5°, rather than bending counterclockwise to the same degree as the target (see Figure 10). The choice of 5° as the critical value for categorizing wiggle is arbitrary. The quantity 2 π was added when the angle reached 360 degrees to prevent the angles from resetting at zero at the base of the loop.

Figure 9. Three tangent lines (red) on a perfectly drawn target loop demonstrate that angles of tangent lines continually increase around the loop when the loop is drawn counterclockwise.

The measure of wiggle was obtained by using the secant to approximate the tangent at each point on the participant's curve, where $\underline{r_t}$ represents the approximation to the tangent at a point, *p*, on the curve. The approximation was obtained by computing the secant,

$$r_t = r_a - r_b \qquad \text{(Formula 3)}$$

where r_a is the vector to the point on the curve that was sampled ten sampling intervals *after* point *p*, and r_b is the vector to the point on the curve that was sampled ten sampling intervals *before* point *p*. Since sampling intervals were approximately 0.0055 seconds (typically 175-185 hertz, although they ranged from 170-206 hertz at times), drawings at constant speed generated 7500-9000 sampling points, while drawings at changing speeds generated 5000-7000 sampling points, depending on the average speed of the writer.

A *Mathematica*® program was written to estimate the tangent at each point on the loop sampled, on average, at 187 hertz (175-200 Hz) by participants, by calculating the angle of the secants. The ArcTan function converted the slopes of these tangents into a list of angles. Since angles in a circular motion proceed from 0 to 2π radians and repeat, all curves will demonstrate increases in angles, followed by sudden decreases as the angle crosses 2π radians. Multiples of 2π were added at those points to obtain a continuous angular variation. The amount of wiggle in a writing line was determined by calculating the percentage of secants over the sample for which the angle decreases, indicating percentage of clockwise motion over the target curves. The *Mathematica*®

program output included the percentage of 21-point sampling units in which the angle of the trajectory was decreasing by more than 5° (moving away from the direction of the target trajectory) and showed the location of the *wiggle* segments on the sample, as seen in the sample shown in Figure 10.

Clockwise curves (wiggles) 9.91 % (Threshold factor: 0.1013)

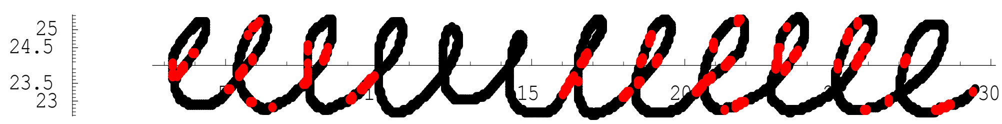

Figure 10. Percentage and location of wiggle in a participant's recorded trace.

The threshold factor (0.1013) registered by the *Mathematica* ® subroutine is attributable to noise from the electronic digitizing tablet and is subtracted from the percentage of clockwise 21-point sampling unit segments that had angles that decreased by more than 5° (i.e., moving clockwise instead of counterclockwise) along the recorded trace. Thus, the true percentage of 21-point sampling unit segments with decreasing angles in this sample is 9.9100 - .1013 = 9.8087%.

3.7. Statistical Analysis

In the present study, generalized additive mixed effect models were used to analyze the timing error and spatial error data and a linear mixed effects model was used to analyze wiggle data. Gender was included in these models to obtain a preliminary estimate of the differences in outcomes by females and males. Because of unequal and small sample sizes for males (n = 12) and for females (n = 23), any findings regarding the effect of gender on timing errors, spatial errors, or wiggle were considered exploratory due to an underpowered analysis. In Chapter 4, *Statistical Analysis*, the analysis for each outcome variable is described and interpreted in detail.

CHAPTER IV: STATISTICAL ANALYSIS

4.1 Introduction

The aim of a statistical analysis is to organize and summarize information from sample data to explain underlying relationships among the variables. In many situations, the goal of the researcher is to extend the findings from a sample of individuals to a larger population that shares the same characteristics (Coolidge, 2000; Devore & Peck, 2000), allowing us to explain and predict the behavior of the population (Faraway, 2006). Researchers in the behavioral sciences often use linear regression models or analysis of variance (ANOVA) models for these tasks. With recent statistical research and advances in computing, new statistical methods have been developed that are more sensitive to data and flexible in addressing analytical issues that have posed problems with traditional methods. In the present study, several recently developed approaches were used, namely linear mixed effects models (LMMs) and generalized additive mixed models (GAMMs), to handle features of the data that could not be adequately addressed by traditional models.

In the sections to follow, I will first give an overview of the goals of the statistical analysis and the assumptions that underlie traditional tests for analyzing data. Second, I will describe features of linear regression, and extend them to linear mixed effects models (LMEs), generalized additive models (GAMs), and generalized additive mixed models (GAMMs). Then I will describe the responses analyzed in the present study (i.e., wiggle, timing errors, and spatial errors). I will begin each discussion with data exploration plots, displaying the graphical features of the data that indicated that the statistical assumptions required for regression were violated. Then, I will explain how LMEs and GAMMS were found to be more appropriate for the data in the present study.

The use of linear mixed models and generalized additive mixed models to analyze the observations in this study can be justified by procedures outlined in several well-regarded statistical texts, but my statistical decisions may legitimately be challenged. Since these models are "on the frontier of statistical research" (Zuur, Ieno, Walker, Saveliev, & Smith. 2009), various statisticians have proposed alternate approaches for analyzing data with these characteristics, with sometimes differing results. For example, the statistical program *R* used in this study contains four packages for producing mixed models based on different statistical theories that vary in methods and interpretation. Thus, I offer my approaches in this dissertation with the knowledge that future research developments in this area of statistics may favor different approaches and results.

4.1.1. Use of the Statistical Program, R

Data in the current study were analyzed with the statistical package *R* (R Development Core Team, 2008) using the lme (linear mixed effects) function from the *nlme* (non-linear mixed effects) package (Pinheiro, Bates, DebRoy, Sarkar, & the R Core Team, 2008) for linear mixed effects models and the package *mgcv* (Wood, 2004; 2006) which uses the generalized cross validation criterion or an unbiased risk estimator (UBRE) for generalized additive mixed effects models (Crawley, 2007).

Parameter estimates, statistical tests, and significance levels for the significant explanatory factors are shown in their respective sections in this chapter and in Chapter 5, *Results*. The developers of the *nlme* and *mgcv* packages in *R* have chosen the Wald test to assess the fixed factors in linear mixed models and generalized additive mixed models. The algorithms in these packages take into account the fixed effects, random effects, and correlated data and assess the null hypothesis that the estimate of the fixed factor ($\hat{\beta}$) equals a hypothesized value (β_0) versus the alternate hypothesis that the estimate does not equal the hypothesized value, i.e., H_0: $\hat{\beta} = \beta_0$ vs. H_A: $\hat{\beta} \neq \beta_0$ with a t-test. The value, *T,* is calculated by Formula 4 (Wood, 2006, p. 8) and follows an approximate t-distribution on n-1 degrees of freedom (West et al., 2007).

$$T \equiv \frac{\hat{\beta} - \beta_0}{\hat{\sigma}_{\hat{\beta}}} \sim t_{n-1} \qquad \text{Formula 4}$$

The p-value, as shown by Formula 5 (Wood, 2006, p. 8), is the probability that the value for *T* would be as extreme as the critical value, *t,* if the null hypothesis were true.

$$p = \Pr\ |T| > |t| \qquad \text{Formula 5}$$

Thus, instead of *F*-statistics that are typically found in ANOVA output for fixed effect models, the data analyses will be expressed in terms of *t*-values with their associated *p*-values. Comparisons of models with random factors will be performed with likelihood ratio tests that will be explained with the descriptions of the models used to analyze the predictors.

4.1.2. Goals of the Statistical Analysis

The statistical analyses in the present study were used to evaluate the smoothness of the recorded traces of the participants' writings and to assess their timing and spatial errors when tracking a moving target. Three response variables were analyzed: (a) spatial accuracy errors in reproducing the shape and size of the target; (b) timing errors, reflected by differences between the times that the participants and targets reached landmarks in the sample; and (c) wiggle, which reflected the lack of smoothness in the contours of the participant writing.

Participant responses were observed under four experimental conditions:

(1) The *Feedback:Present-TargetRate:Constant* condition, where the participant's trace was displayed on the screen concurrently with the target which moved across the target template at a constant rate.

(2) The *Feedback:Present-TargetRate:Changing* condition, where the participant's trace was shown on the screen concurrently with the target, which abruptly changed rate four times in the middle of the target template.

(3) The *Feedback:Absent-TargetRate:Constant* condition, where the participant's trace did not appear on the screen during drawing, while the target moved at a constant rate over the target template.

(4) The Feedback:Absent-*TargetRate:Changing* condition, where the participant's trace did not appear on the screen during drawing, while the target abruptly changed rate four times in the middle of the target template.

Drawing speed was not included in these analyses, on the advice of Gary Oehlert, Professor, University of Minnesota School of Statistics (personal communication, June 11, 2008), because drawing speed was not completely under the control of the researcher and could operate as both a predictor and a response variable. Further, although participants were instructed to couple their motions with the speed of the targets, it was not clear how they used the information about the speed of the target. For example, they may have attempted to match the target speed, ignored it in favor of increasing spatial accuracy, failed to detect it, or chosen a solution than was compromise between competing timing and spatial goals. Thus, it was not possible to determine if drawing speed functioned as a predictor, and outcome, or somewhere in between those two options. These possibilities produce a situation where a variable may serve alternately as a predictor and as an outcome, resulting in a correlation between a predictor and error. The formal name for this occurrence is termed *endogeneity* and is problematic because it obscures a clear understanding of the effect of the predictors in the model (Wooldridge, 2008), by producing biased and inconsistent estimates and misleading results (Shepherd, 2008). Variables that induce endogeneity should be avoided if possible, and require very special handling when they do occur. To prevent this problem, drawing speed was not assessed for its effects on wiggle, timing errors, or spatial errors across all the experimental conditions in the present study.

4.1.3. Assumptions of Statistical Tests

Statistical models rely on conditions about the data, termed *assumptions,* which must be met so that reliable interpretation of the results can be made. Statistical inference, the process of making generalizations and predictions from sample data, is valid only if these assumptions are met.

Traditional models that are based upon the linear model framework rely on a set of four assumptions:

(1) The errors are independent.

(2) The errors are normally distributed with mean 0 and variance σ^2.

(3) The residuals are homogeneous (the same spread).

(4) There are no patterns in the residuals, i.e., the variance of the residuals is not a function of the predictors, producing non-constant variance, also known as *heteroscedasticity* (Oehlert, 2000, pp. 111-124).

Assumptions are tested through graphical and numerical assessments (Oehlert, 2000). The linear model may be robust against minor violations of normality in the residuals, but more severe violations require an alternative statistical approach to traditional ANOVA or regression-based methods. Assessing assumptions is therefore necessary to prevent overestimating or underestimating the standard errors, which would lead to errors in the inferences made from the analysis. This step should not be overlooked: Zuur, Ieno, and Smith (2007) described an example where a traditional linear model was applied to correlated data that violated the assumption of independence, and as a result, "the *t*-values obtained by linear regression were inflated by approximately 400%!" (p. 279). Thus, an important aspect of data exploration involves verifying that the assumptions of the chosen statistical approach are met. When assumptions are violated, a different statistical model must be chosen that minimizes the effects of those violations or accommodates them. This step is especially important with behavioral data, which are often messy and violate assumptions of standard statistical tests.

4.1.3.1. The Problem of Dependent Data

The first assumption, independence, is identified from understanding the nature of the data and the system that was used to collect it. “Independence is the most important of these assumptions and the hardest to accommodate when it fails” (Oehlert, 2000, pp. 111-112). Observations taken several times on the same participant (i.e., repeated measures), violate the fundamental assumption of independence of observations since the “peculiarities of an individual are reflected in all the measurements" taken on the individual (Crawley, 2005, p. 13). In the case of repeated measures, where some data are known to violate the independence assumption, a more advanced model (e.g. a linear mixed effects model) may be chosen *a priori* to analyse the data.

4.1.3.2. The Problem of Non-Constant Variance

For some models, the residuals show a pattern, violating the assumption of constant variance. Non-constant variance, or heteroscedasticity, means that the variance of the residuals changes as the values of the predictor(s) vary and suggest that the model is misspecified, providing a poor fit to the data (Oehlert, 2000; Zuur et al., 2007). When heteroscedasticity is present in a model, standard errors and significance tests (p-values, t-values, and F-values) will not be accurate (Cohen, Cohen, West, & Aiken, 2002; Keith, 2006; Oehlert, 2000).

When heteroscedasticity is found in the residuals, the analyst has a wide array of options for dealing with this problem. In general linear models, such as ANOVA and linear regression, non-linear relationships between explanatory and response variables can sometimes be handled with transformations, or by adding exponential terms, interactions, or smoothing methods. These strategies are intended to protect against invalid inferences from inaccurate standard errors (see Oehlert, 2000, for a review). Zuur et al. (2007) noted that transformations and interactions are not always effective when nonlinearity is present. The patterns of residuals should be carefully examined before applying a transformation or adding terms to a model with non-constant variance, since the patterns in the residuals may suggest important features of the subject under study (Carroll, 2002). When the residuals suggest nonlinearity that may be useful to understand, smoothing techniques and methods for imposing covariance structures on the residuals may provide alternative methods for decreasing non-constant variance in data sets consisting of repeated measures.

In the sections to follow, I will motivate the discussion of more advanced models by outlining the linear regression modelling framework and showing how it was extended to create a linear mixed effects model that was used to analyze the wiggle data. Following the description of the linear mixed effects model, I will extend this approach further by describing the generalized additive mixed model (GAMM) that was used to analyze the spatial and timing error data. Finally, I will show how covariance structures in the GAMM for the spatial data were adjusted to deal with non-constant variance in present study.

4.1.3.3. The Starting Point: Data Exploration with Graphical Displays

The first step in statistical analysis is data exploration through plots and graphical displays. Diagrams of the data structure illustrate relationships among variables that may not be apparent in numerical summaries. Plots are essential and efficient methods for determining central tendencies, spread, and distribution of the data (Zuur et al., 2007). Fox (2002) reported that "important characteristics of data are often disguised by numerical summaries and – worse – the summaries can be fundamentally misleading. […] Statistical graphs are central to effective data analysis, both in the early stages of an investigation and in statistical modelling" (p. 27).

Exploratory plots are especially useful for fostering logical model development by determining the structure of the data, identifying correlated observations, and exploring relationships within the data. Independence is tested “by employing a series of plotting methods involving the residuals, the fitted values, and the predicted random effects" (Crawley, 2007, p. 628) of linear models. In the model development phase, patterns of residuals are assessed to determine whether the models meet the necessary assumptions or if corrective measures are necessary to make valid and reliable inferences from our models (Oehlert, 2000). Although behavioral sciences traditionally have employed tests of sphericity such as Mauchly's test (Field, 2005) for assessing nonconstant variance, Draper and Smith (1998) reported that residual plots are more informative than numerical tests for this assessment. Oehlert (2000, p. 118) is more emphatic ("*do not use them!*"), observing that these tests are highly sensitive to non-normality. Thus, if they provide usable information at all, these tests do not inform us what steps are needed to address non-constant variance when it is found. Like Draper and Smith (1988), Oehlert (2000) advises the use of exploratory plots to evaluate whether the data reasonably meet the assumptions for the selected statistical tests and models. Thus I chose to begin an assessment of relationships in the data through diagrams and exploratory plots in the present study. Diagrams and plots for wiggle data will be found in Sections 4.4.3 and 4.4.4; for timing errors in Sections 4.6.3 and 4.6.4; and for spatial errors in Sections 4.8.3 and 4.8.4.

4.2 Introduction to Mixed Effects Models

4.2.1. Inadequacies of Traditional Models for Studies Using Repeated Measures

Repeated observations from the same participant or unit of analysis present a challenge for statistical analysis because repeated measures form a natural set of correlated errors. Correlation among errors violates the most important assumption of the linear regression model, namely independence (Oehlert, 2000). When statistical assumptions are violated, standard errors are often decreased, and estimates of significance are inflated (Chatterjee & Hadi, 2006, Luke, 2004; Zuur et al., 2007), making traditional statistical tests unreliable.

Several methods for dealing with repeated measures are reported in the behavioral literature. One traditional approach resolves the problem by averaging the repeated measures for each participant, then using the averages in a fixed effects model. The averaging approach is not preferred, because it is likely to bias the results (de Leeuw & Meijer, 2007; McCall & Appelbaum, 1973), inflating *F*-values and leading to rejection of the null hypothesis when it is true. Another approach for analyzing multiple observations on a single experimental unit is to use repeated measures ANOVA, which is applicable when errors are normally distributed, variances of each condition are equal, and covariances between all variables are equal (Heath, 2000). However, behavioral data rarely meet these requirements. Even if these stringent requirements are met, repeated measures ANOVA procedures require balanced data, meaning that cases with missing data must be thrown out. Although classic repeated measures is a linear mixed effects models (Oehlert, personal communication, March 9, 2010), Baayen, Davidson, & Bates (2008) observed that the classic approaches for analyzing repeated measures data suffer multiple drawbacks, which can be eliminated by using a more complex model. These drawbacks include: (a) deficiencies in statistical power related to the problems posed by repeated observations; (b) lack of a flexible method of dealing with missing data; (c) disparate methods for treating continuous and categorical responses; and (d) unprincipled methods of modeling heteroscedasticity and non-spherical error variance for either participants or items. Thus, traditional approaches rarely fit the type of repeated measures data found in behavioral sciences, and may be costly in terms of decreased power and reduced information available for analysis.

4.2.2. Features of Mixed Effects Models

Over the past three decades, the problem of correlated errors has generated considerable statistical research and a variety of approaches for dealing with them, including mixed effects models (Faraway, 2006; Pinheiro & Bates, 2004). Like fixed effect models, mixed effects models belong to a common family of procedures that are based on the general linear model, including ANOVA and regression (Keith, 2006; Nelder & Wedderburn, 1972). Mixed effects models are specifically designed for use with data that are grouped or nested (i.e., arranged in hierarchies) where observations may be more highly correlated with some observations than with others. According to Pinheiro and Bates (2004, p. 3), “mixed-effect models are primarily used to describe relationships between a response variable and covariates in data that are grouped according to one of more classification factors." West et al. (2007) suggested that analyzing grouped or hierarchical (nested) data add flexibility to a model, making it possible to compare data on more levels than is possible in linear regression.

Laird and Ware (1982) were among the first to develop standard notation for mixed effects models, which popularized their use for analyzing special features of repeated measures data in many discipline. In education and social sciences, mixed effects models have been structured as hierarchical linear models, or HLM (Bryk & Raudensbush, 1992; Raudenbush & Bryk, 2002), multilevel models (Kreft, & de Leeuw, 1998), random coefficient models (Longford, 1993), and split-plot designs (Edwards, 1984; Oehlert, 2000). Other mixed effects models include time series and spatial models. The mixed effects model framework consists of fixed and random covariates (i.e., predictors or explanatory variables) that produce an effect on the response. By including both fixed and random variables, mixed models allow for estimation of random variation between groups and within groups (Gelway, 2006, p. 1), while preserving power and controlling bias.

A mixed effects model is expressed as

$$Y_i = \text{fixed variable(s)} + \text{random variable(s)} \qquad \text{(Equation 1)}$$

where an observed response (Y_i) is predicted from the fixed and random variables.

Since texts on mixed effects models do not provide uniform definitions of fixed and random variables (Gelman, 2005), the following definitions from Searle, Casella, and McCulloch (1992) are used in this dissertation:

(1) *Fixed effects* consist of one or more continuous variables (covariates) or categorical factors with a finite number of levels whose effects on the outcome are of specific interest. Temperature, diagnoses, treatment methods, and Gender are typical examples.

(2) *Random effects* consist of one or more variables representing levels of a factor (e.g., individuals in the present study) that are randomly drawn from a broad population with an nearly infinite number of levels, with the goal of making inferences about the larger population. A group of students selected from the kinesiology classes at the University of Minnesota is a random variable, if the responses of students in general are of interest, rather than the responses of the specific thirty-five students chosen for the present study.

Random effects in linear mixed models may be structured with random intercepts where the means within the groups are allowed to vary, called *random intercept* models; random slopes where the rates of change are allowed to vary, called *random slope* models; and random intercept and slope models where both the means within the group and the rates of change are allowed to vary, called *random coefficient* models. A range of structures have been developed for the covariances in these models which may be applied to all participants or to subgroups in the data (Zuur et al., 2009). By including both fixed and random variables, mixed models allow for estimation of random variation between groups and within groups (Gelway, 2006), while preserving power and controlling bias.

More recently, mixed effects models have been developed for nonlinear data. Nonlinear mixed effects models include generalized linear mixed models (Breslow & Clayton, 1993; Pinheiro & Bates, 2000), additive mixed effects models (Wang, 1998; Verbyla, Cullis, Kenward, & Welham, 1999; Zuur et al., 2007), and generalized additive mixed effects models (Lin & Zhang, 1999). These models will be described and applied to the present study later in this chapter.

4.2.2.1. Advantages of Mixed Effects Models: Generalizability

The results of a model with only fixed predictors, such as traditional ANOVA and regression, show how the explanatory variables affect a particular sample by influencing only the mean of Y_i (Crawley, 2007, p. 627). However such fixed effects models are limited, because they address only "the specific sample... used in the experiment, while the main interest is the population... from which the sample is drawn. In particular, [fixed effects models] do not provide an estimate of the between-[subject] variability...“ (Pinheiro & Bates, 2004, p. 7). Wood (2006) concurred, noting that a fixed linear effects model does not permit predictions about the population of individuals that are represented by the participants, even though we expect that members of the population will generally respond in the same way as the study participants. Wood (2006) explained that fixed models are problematic for repeated measures data because (a) the population mean (α) and coefficients for the fixed terms (β ’s) are confounded in repeated measure analyses and (b) fixed effects models provide information only on the units under study. Thus, when participants are modeled as a fixed effect, the results cannot be used to predict the behavior of other individuals in the population. Pinheiro and Bates (2004) concluded that a random-effects model circumvents these problems by treating the subject effects as random variations around a population mean. Crawley (2007, p. 627) concurred, explaining that random effects influence only the variance of Y_i , enabling us to understand how the effects of the predictors vary for different experimental units or participants in the study. Thus, modelling participants as random effects makes it possible to explore differences among individual participants by calculating measures of within-subject and between-subject variability, allowing inferences to populations who share their characteristics.

In the present study, the participants consisted of thirty-five healthy male and female college students, randomly recruited from a much larger, apparently homogeneous group of students taking kinesiology classes at the University of Minnesota. Since random selection from the population allowed replacement of any participant by another member of that population, participant responses could be used to predict the performance of similar populations of healthy college students (Snijders & Bosker, 1999).

4.2.2.2. Advantages of Mixed Effects Models: Management of Missing Data

Missing data pose serious problems for traditional analyses based on balanced data, such as *t*-tests and standard or repeated measure ANOVAs that use ordinary least square solutions. Unlike traditional approaches, mixed effects models use Restricted Maximum Likelihood (REML) solutions, which can handle unbalanced groupings of data (Patterson & Thompson, 1971). According to Howell (2008), REML methods are not constrained by the assumption of balanced data, preserving more data for analysis when cases are missing at random (i.e., when the missing data are not produced by a systematic pattern and they share the same characteristics as observed data).

4.2.2.3. Advantages of Mixed Effects Models: Estimating Within-Subject Variability

Mixed effect models make it possible to estimate the variability among scores produced by a participant (within-subject variability) and across participants (between-subject variability). Variability is estimated by the intraclass correlation (also called the intraclass correlation coefficient, ICC), which refers to the proportion of variance in the unit of analysis (i.e., person, group, set of items, etc.) compared to the total variance in the data (Hox, 2002). When participants are modelled as random effects, the mixed effect model imposes a correlation structure between dependent observations. Since random effects influence only the variance of the outcome variable (Crawley, 2007), intraclass correlations allow us to estimate the relative contributions of within-subject and between-subject variability. The correlations between observations from the same participant, also called the compound correlation (Faraway, 2006, p. 170; Pinheiro & Bates, 2004), are used to calculate the intraclass correlation,). Intraclass correlations provide a numerical value for the relationship of the variation within subgroups to the variation represented in the data set (Singer & Willett, 2003), giving an estimate of similarities within members of a group. The general formula for intraclass correlation is given by

$$\mathrm{ICC}_{\boldsymbol{a}} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2} \qquad \text{(Formula 6)}$$

Please note that this formula can be expanded to compare more levels of variances in more complex models.

In Formula 6, the term σ_a^2 represents the variance for the subgrouping whose observations are being compared, while σ^2 represents the variance for the residuals in the model. According to Hox (2002), the intraclass correlation is the "proportion of group level variance compared to the total variance… [i.e.,] the expected correlation between two randomly chosen units that are in the same group" (p. 15). For example, if scores from each participant are the subgroup of interest, a high intraclass correlation would result from large variations between separate individuals (between-individual variance, or σ_a^2), which produce a high value for σ_a^2 relative to σ^2 . High intraclass correlations suggest that scores within participant subgroupings are more similar to each other (i.e., more highly correlated) than to observations from the group as a whole. High intraclass correlations are characteristic of studies where individuals perform consistently over their own trials and quite differently from one another, giving consistently higher (or lower) scores than the average score for the all of the participants in the study. A low intraclass correlation would indicate that the variation in the data set is evenly distributed across the subgroups, indicating that the subgroups (or set of scores from individual participants, in this case) produced approximately the same scores as the others. Intraclass correlations are important when choosing a model because standard linear models assume that the data are independent and normally distributed, which would be reflected by low intraclass correlations. Since linear regression ignores high intraclass correlations, it is an inappropriate model for data with highly correlated observations.

4.2.3. Application of Mixed Effects Models in the Present Study

The current study involves analysis of repeated measures for each participant. Thus, observations were likely to be correlated. In this case, mixed effects models may be better suited for analysis of the data than traditional methods, for reasons described previously. The fixed effects for wiggle data included the categorical explanatory variables, Feedback (present and absent), Target Rate (changing and constant), and Gender (female and male). In addition to these variables, terms for peak location (2-12) and peak-to-peak segments (2-12) were added as fixed effects to assess changes along the trajectory for timing errors and spatial errors, respectively. Random effects included sets of scores contributed from repeated measures of participants.

4.3. Building Mixed Effects Models

Despite the broadening acceptance of mixed model statistical analysis, some researchers are reluctant to use these methods because there is no single approach or best way to construct a multilevel model. Much depends on the hierarchical relationship of the data and the research questions of the investigator. The following sections describe different structures for mixed models, beginning with a simple linear mixed effects model and its relationship to a linear regression model. First, I will show how the linear regression model can be extended into a simple linear effects model with fixed and random parts with nested data, which was used to analyze the wiggle data in the present study. In this section, I will describe the components of the most basic linear mixed model by building the simplest mixed model framework for nested data, involving multiple observations from an individual. In this example, I will include an intercept for the data set and a random intercept for each participant, assuming the same rate of change for each participant. Models of this type are common when individuals respond in the same way to a treatment, but begin with different characteristics. Following this discussion of random intercept linear mixed effect models, I will briefly describe random slope models and models that combine random intercepts and random slopes. Models that are more complex are common, and are discussed in Crawley (2007), Faraway (2006), Galwey (2006), Pinheiro and Bates, (2004), West et al., (2006), Wood (2006), and Zuur et al., (2009), among others. Then I will expand the linear mixed effects model to generalized additive mixed effects models, which reduce nonlinearity in data by using smoothing techniques. Following these descriptions, I will explain how this type of model was used to analyze nonlinear timing and spatial error data in the present study.

4.3.1. Linear Regression Modeling

In its most elementary form, a linear regression model can be written as:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad \text{(Equation 2)}$$

where the variable Y_i is the value of the response variable in the *i*th observation. X_i is a known variable that represents the value of the independent variable in the *i*th observation (Neter, Wasserman, & Kutner, 2000). The symbols α and β stand for unknown regression parameters, with α representing the *Y*-intercept of the regression line for the model, a constant denoting the mean of the response if $X_i = 0$, and β representing the "slope of the regression line, the probability distribution of *Y* per unit increase in X_i" (Neter et al., 2000, p. 33). Weisberg (2005) has described β as the slope of the best-fitting line denoting "the predicted value of the response when the predictor is fixed at the value $X = x$" (p. 8). The term ε_i represents the errors, the difference between the true values, Y_i, and the values of Y predicted by the model for a particular value of *X* (Verzani, 2005) where the mean expectation for the residuals is 0 and the variance is σ^2. Residuals can also be described as deviations from predicted values, i.e., information in the sample that is not fully explained by the model (Field, 2005). Since Equation 2 depicts a traditional fixed effect model, no random terms are included.

Figure 11 represents a plot for a simple linear regression model, with the bold line reflecting the expected response of the population. Please note that the data used for Figures 11 – 14 were created for illustration purposes; they were not taken from the present study.

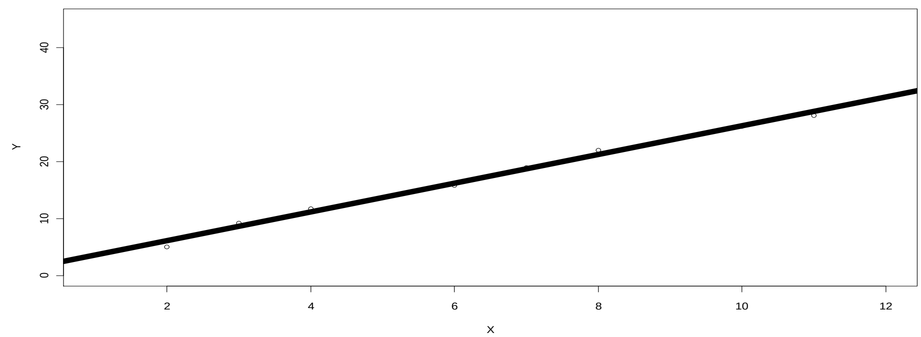

Figure 11. Linear regression plot with illustrative data.

Figure 11 depicts a statistical relationship between an explanatory fixed variable represented by βX_i, and an outcome represented by Y_i. The small points that deviate from the line (i.e., residuals) are represented by ε_i, which are distributed with mean 0 and variance σ^2. In mathematical notation, we write $\varepsilon_i \sim N(0,\sigma^2)$.

4.3.2. Basic Linear Mixed Model Structures

As described previously, the simplest mixed effects model can be written as

$$Y_i = \text{fixed variable(s)} + \text{random variable(s)} \qquad \text{(Equation 1)}$$

where both fixed effects and random effects function as predictors of the response. As discussed in section 4.2.2, I am interested in the effects of specific levels of the factors Feedback (absent and present), Target Rate (changing and constant), Gender (Female and Male), and the order of the peak or peak-to-peak segment on the response, rendering these variables *fixed effects.* Since participants could be considered to be drawn randomly from a large group of similar individuals (i.e., University of Minnesota students), participants are modelled as random effects with the goal of making inferences to a larger population of similar individuals.

Mathematically, a simple linear mixed effects model with fixed and random terms is depicted by

$$Y_{ij} = \alpha + a_i + \varepsilon_{ij} \qquad \text{(Equation 3)}$$

One immediate difference between traditional models and linear mixed models is that traditional models rely on a single index *i* for all observations, where linear mixed models index each observation in a hierarchical or nested system in which one unit of analysis is made up of subunits that are uniquely assigned to it. In this example, each observation is represented by a two- index system, *ij*, which stands for observation *j* from participant *i* in 2-way nested data sets where observations are nested in participants. This indexing system can be expanded further to include additional levels if the data have added layers of structure.

Arrangement of data into subgroupings in mixed effect models allows estimation of within-subject variation and between-subjects variation for each subject. The price of this arrangement is only a few parameters, depending on the structures of the data. For example, in the case of a simple random intercept model, two parameters are used, namely σ_a^2 for random intercept models with individual-specific means and σ_b^2 for repeated observations for each individual, making this model very efficient (Pinheiro & Bates, 2004, p. 8). By restricting within-subject variation and between-subjects variation for each subject with a linear mixed model strategy that uses only σ_a^2 and σ_b^2, it is possible to avoid pseudoreplication, inflated degrees of freedom, and spurious p-values in repeated measures studies (Crawley, 2007; Zuur et al., 2007). Thus, this structure spares large number of degrees of freedom to specify participant effects, increasing the power of the model, and ensuring the reliability of the statistical tests.

4.3.2.1. Random Intercept Mixed Effects Models

The random intercept mixed effects model illustrated in this section provides an average response overall (α) and an average response for each participant (a_i). Equation 3 presents a linear mixed model in which the fixed part contains only the intercept, α, and a random intercept a_i for each participant. This simple framework model asks whether there are any differences in the mean response across participants (a_i) and assumes that the responses are independent and identically distributed, with homogeneity of residuals (i.e., with no patterns in the residuals plot, described as *iid,* or independent and identically distributed).

Referring again to Equation 3, I will discuss the subscripts below.

$$Y_{ij} = \alpha + a_i + b_i X + \varepsilon_{ij} \qquad \text{(Equation 3)}$$

The term a_i is an intercept for the random term specific for each participant *i,* denoted by the subscript *i*. In the present data set, the index *i* runs from 1 to 35 (i_1-i_{35}) to represent each of the 35 participants. The term a_i is assumed to be normally distributed with mean 0 and variance σ_a^2. The term b_i allows for participant-specific slopes. Thus, the model in Equation 3 suggests that the outcome Y_{ij}, a specific observation, is defined by four elements: the expected population intercept α, the expected participant-specific changes a_i from the expected population intercept α, the slope b_i for each participant, and errors ε_{ij}. The errors, represented by ε_{ij}, are associated with each individual observation within participant *i,* and represent deviations from the expected participant-specific intercept with which they are associated. These errors are normally distributed with mean 0 and variance σ^2.

This structure is depicted by Figure 12. Recalling that Figure 11 in Section 4.31 represented a traditional linear fixed effects model, i.e., the expected mean value at each landmark in the sample, note that individual participants are represented by different colored lines in Figure 12, each denoting a random change from the population intercept for the overall average. This plot presents each set of participant responses with the same slope, meaning that the participants respond similarly to the explanatory variables over multiple measures, i.e., the points (x = 1-12), as shown by identical slopes. Figure 12 adjusts for individual differences at the start of the treatment, shown by differences in the intercepts.

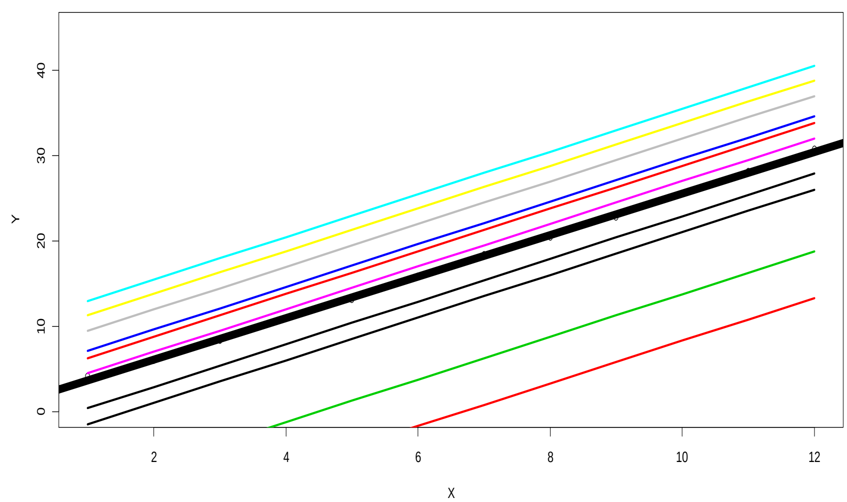

Figure 12. Linear mixed model for ten fictional participants, with the mean scores at each landmark (x) designated by the black line and a separate colored line designating the mean scores at each landmark (x) for each participant. The colored lines

for each participant assume that the participant provided observations over several trials to yield scores for the means.

4.3.2.2. Expanding the Random Intercept Linear Mixed Model to a Three-Way Data Set

The linear mixed model can be expanded beyond the two-level data sets described in Section 4.3.2.1 to include additional correlated data. In this example, I will explain a three-level data hierarchy with sample observations nested in participants.

The three-way nested model, denoted by Equation 4, allows each participant a_i to have a different intercept and allows for differences among trials performed by each participant, but imposes the same effects (i.e., same slope).

$$Y_{ijk} = \alpha + a_i + b_{j|i} + \varepsilon_{ijk} \qquad \text{(Equation 4)}$$

The difference between the two-way random intercept model denoted by Equation 3, and the three-way random intercept model with nested data in Equation 4

is the addition of the term $b_{j|i}$ and the expansion of the residual term from ε_{ij} to ε_{ijk}. In a three-way nested data set, each observation is represented by this three-index system. The subscript $_{ijk}$ stands for observation *k* in sample *j* from participant *i,* and assumes that several observations were taken from each sample. Although this notation may seem complicated at first, there are few differences between the three-way random effect notation and single-index systems denoted by the two-way random intercept model in Equation 3. First, the three-way notation allows each participant to have unique mean values which may be higher or lower than for other participants (a_i) and allows each sample nested in each participant ($b_{j|i}$) to have higher or lower values than the mean for the participant. In the present data set, the term a_i is an intercept specific for each participant (i_1-i_{35}). The notation $b_{j|i}$ refers to an intercept for each trial$_j$ (j_1-j_{16}), nested within the corresponding participant, making this model a three-way structure.

Thus, the model in Equation 4 suggests that the outcome Y_{ijk}, a specific observation, is defined by four components: (a) the average response (α); (b) the term a_i denoting intercepts running from 1-35 to represent each participant$_i$, in the study (i_1-i_{35}); (c) the term $b_{j|i}$ representing each trial j ($_{j1}$-$_{j16}$) which runs from 1-16 to represent the sixteen trials that are nested within the corresponding participant$_i$; and (d) the residuals, denoted by ε_{ijk}, signifying the deviations of individual scores from the mean for the trial at a particular point. ε_{ijk} represents variation that is not accounted for by the model. Both terms a_i and $b_{j|I}$ are assumed to be normally distributed with mean 0 and variance σ_a^2 and σ_b^2, respectively. The random effects a_i, $b_{j|i}$, and ε_{ijk} are assumed to be independent.

The three-way nested random effects model is depicted by Figure 13. As in Figures 11 and 12, the solid black line depicts the overall group mean at each sampling point (x) without individual-specific distinctions. The position of each individual-specific intercept a_i is denoted by a solid colored line. The average observations for individual-specific trials are nested in participants ($b_{j|i}$), and denoted by colored dotted lines centered on the same- color individual-specific lines.

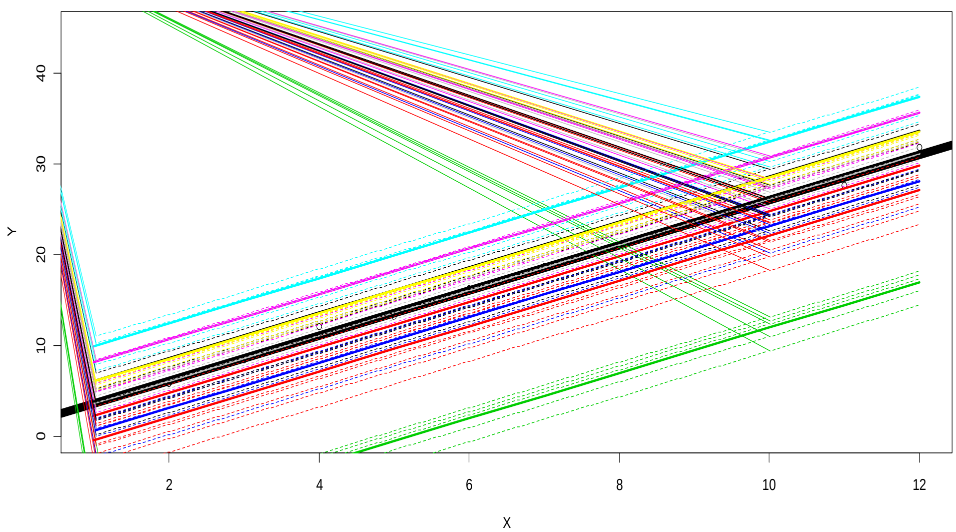

Figure 13. Lines reflecting the hierarchical structure of a three-way nested data set in a random intercept mixed model.

4.3.2.3. Random Slope Mixed Effects Models

In some repeated measures studies, explanatory variables have varying effects on each participant, requiring a more complex model strategy. The random slope framework provides a mechanism for modeling an individual-specific rate of change over the points in the sample (b_i). Equation 5 presents a linear mixed model in which the fixed part contains only the intercept, α , and a random component consists of slope b_i for each participant. This model suggests that the relationship between the explanatory variables and responses are real, but that they differ for each participant. The equation for the simplest random slope model is

$$Y_{ij} = \alpha + b_i X_j + \varepsilon_{ij} \qquad \text{(Equation 5)}$$

In a random slopes model, study participants begin at the same value, but vary in their responses to the explanatory variables. The model in Equation 5 suggests that the outcome Y_{ij}, a specific observation, is defined by three elements: the group mean, α ; the average participant- specific slope of (b_i), which allows us to visualize participant-specific mean deviations from the average effect (slope) of the explanatory variable; and errors ε_{ij} representing individual observations that deviate from the participant-specific means. The term b_i is assumed to be normally distributed with mean 0 and variance σ_b^2.

A simple random slope model is depicted in Figure 14. The colored lines in Figure 14 reflect participant-specific mean responses measured systematically across the sample. In other words, each line indicates individual-specific relationships between the explanatory variables and their mean response measured sequentially at points on the sample. Figure 14 shows each participant responding differently to the treatments, although they all began at the same level on the variable of interest (single intercept).

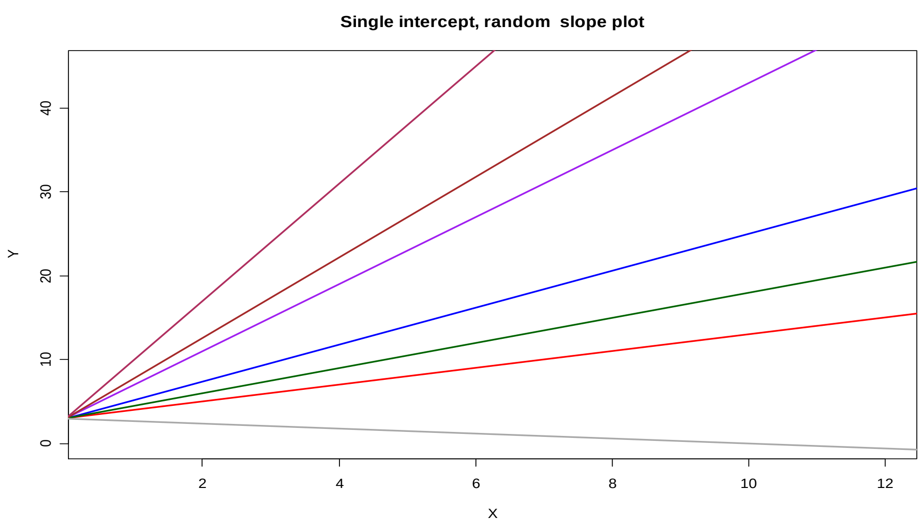

Figure 14. Linear mixed model with individual-specific random slopes for each participant, denoting each participant's average response at each landmark (*x*) in the sample.

4.3.2.4. Random Intercept and Slopes: Random Coefficient Models

Repeated measures studies may require more complex models where individual-specific intercepts and slopes vary. In models with random intercepts and random slopes, individuals are characterized by their respective intercepts and their respective slopes. Snijders and Bosker (1999) note these types of models may be the most realistic and common in analysis of behaviors. Random intercept and slope models account for differences in means and rates of responses between individuals or units of study. For example, a random intercept and random slope model may be used to describe differences between participants on relevant characteristics at the onset of the study (random intercepts) and on their differing responses to the explanatory variables (random slopes). The equation for a basic random intercept and random slope model is

$$Y_{ij} = \alpha + a_i + b_i X_j + \varepsilon_{ij} \qquad \text{(Equation 6)}$$

In this model, each individual (*i*) has a separate intercept a_i and a separate slope b_i at each level of the explanatory variable. The model in Equation 6 suggests that the outcome Y_{ij}, a specific observation, is defined by four elements: the population mean, α; participant-specific changes a_i from the expected population intercept α; the participant-specific changes from the average effect (slope) of the explanatory variable, resulting in a participant-specific slope of (b_i); and errors ε_{ij} which represent individual observations that deviate from the participant-specific intercept and slope means.

The random intercept-random slope (i.e., random coefficient) model is illustrated in Figure 15. The responses for each participant, indicated by different colored lines, denote a random change (a_i; i.e., a random intercept) from the average response overall (α) and a random change (b_i, i.e., a random slope) from the average effect of the explanatory variables on the population.

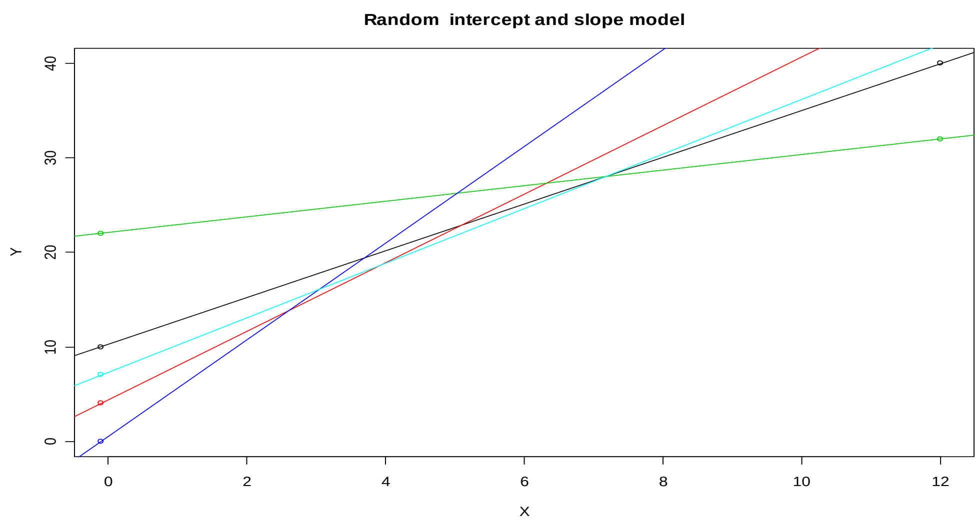

Figure 15. Linear mixed model with random intercepts and slopes for each individual.

4.3.3. Generalized Additive Models and Generalized Additive Mixed Models

I will begin this discussion by describing the generalized additive model (GAM). Second, I will show describe the generalized additive mixed model framework (GAMM). In this brief overview, I will describe smoothers, a feature of GAMs and GAMMs for coping with nonlinearity and non-constant variance in the residuals. Finally, I will show how the GAMM approach applies to the timing and spatial errors in the present study in Section 4.7 and 4.9, respectively.

4.3.3.1. Features of Generalized Additive and Generalized Additive Mixed Models

Generalized additive models (GAMs) and generalized additive mixed models (GAMMs) are extensions of linear fixed effect models and linear mixed effect models, respectively. GAMs and GAMMS were developed to handle nonlinearity in the data by incorporating nonparametric methods, such as smoothing, into the models (Crawley, 2007). GAMs and GAMMs apply smoothing techniques that preserve some aspects of nonlinearity where the nonlinearity itself is interesting. These approaches can be found in many sciences where models contain nonlinear covariate effects.

Like fixed effect models, GAMs are not designed to handle dependence in the data, and therefore do not include random effects. In contrast, generalized additive mixed models (GAMMs) are the generalized extension of linear mixed effects models and are useful in handling the effect of nonlinearity and random effects. Coull, Schwartz, and Wand (2001) reported that generalized additive mixed models are “a relatively recent statistical development (e.g. Wang, 1998; Verbyla et al., 1999; Lin & Zhang, 1999), [which] allow for smooth functional relationships, subject-specific effects and time series error structure” (p. 339) in a data set. GAMMs can be very effective when data include correlated observations, missing values, non-linearity, and non-constant variance in the residuals. GAMMs have been applied in disciplines where data are likely to have these features, including geography (Brown, Goovaerts, Burnicki, & Li, 2002), ecology (Guisan, Edwards, & Hastie, 2002), econometrics (Kim & Linton, 2004), health sciences (French & Wand, 2004), and social sciences (Keele, 2008). GAMMs have only recently been applied in psychology, but may have particular relevance in studies where nonlinear change is measured over time, as in education and development (Barnes, Yaffe, Satariano, & Tager, 2003; Iglowstein, Jenni, Molinari, & Largo, 2003).

A variety of smoothing techniques are applied in GAMMS to reduce the effects of random variation. Simonoff (1996) reports that smoothing methods are particularly useful when the underlying shape of the data is unclear or complex, since they provide “the ability to identify potential unexpected structure in the data” (pp. 5-6) and "allow… the data to tell the analyst what the pattern truly is” (p. 3). Simonoff (1996) comments that “smoothing methods can aid in data analysis …by extracting more information from the data than is possible purely nonparametrically, as long as the (weak) assumption of smoothness is reasonable; and by being able to free oneself from the 'parametric straitjacket' of rigid distributional assumptions, thereby providing analyses that are flexible and robust” (p. 8).

Smoothers may be as simple as the familiar normal curve imposed on a histogram, or involve complex choices and procedures. Keele (2008) remarked, "some authors have even described the choices revolving around the smoothers as controversial (Box-Steffensmeier & Jones, 2004) [leading]…some analysts to see the use of smoothers as more akin to art than science" (p. 93). Details of smoothing methods are described in excellent references on generalized additive models and smoothing procedures including Hastie and Tibshirani (1990), Keele (2006, 2008), Ruppert, Wand, and Carroll (2003), Simonoff (1996), Wood (2006), and Zuur et al. (2009). These authors agree that smoothing methods in GAMMs produce a compromise between achieving a close fit of the data and creating a smooth line that reveals general patterns in the data. Thus, in this study, I used the smoothers built into the *mgcv* package in the statistical language *R* (Wood, 2004, 2006) which apply smoothing splines, specifically cubic regression splines with shrinkage (cs), to the data.

The approach taken by Wood (2004, 2006) in the *mgcv* package in *R* puts a moving window around target values and uses piecewise cubic splines to interpolate between target values using cross-validation to select the window width (Verbyla, Cullis,Kenward, & Welham, 1999). The *mgcv* package includes a two-step process of smoothing that applies cubic splines to interpolate between target points, while penalizing irregularity (i.e., wiggle) in the sample data (Wood, 2004). An example is shown in Figure 16.

In Figure 16, a window is placed around the area designated by the small triangle on the *x*-axis of Figure 16 at peak = 5. All observations inside the window are used to obtain an estimated value at the selected landmark. The *mgcv* program selects the window width and moves the window horizontally across the data, interpolating between target points. The resulting smoother is shown by the solid line in Figure 16.

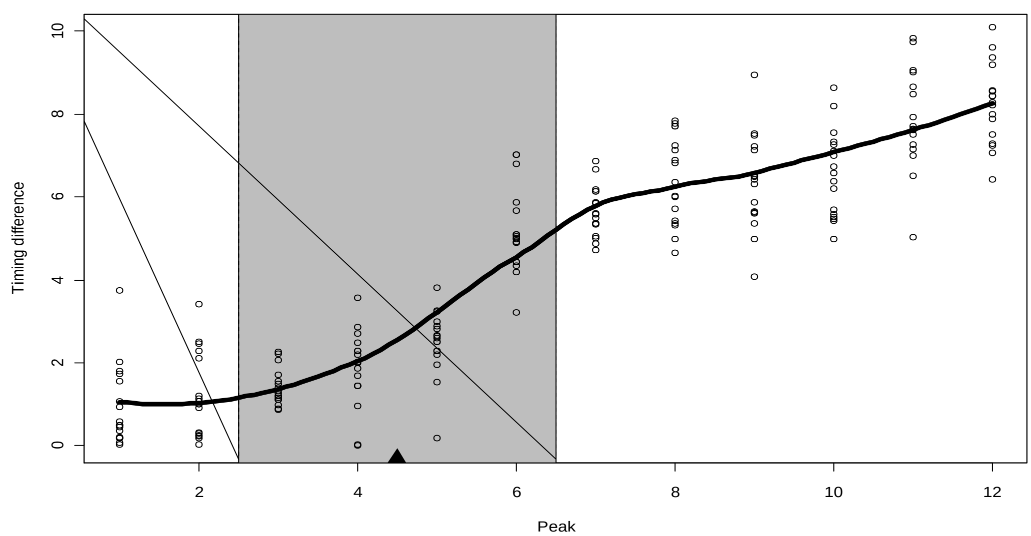

Figure 16. Illustration of a moving window used to calculate a smoother.

Smoother parameters in GAMMs have calibration values indicated by estimated degrees of freedom, (edf; also described as equivalent degrees of freedom by some authors) which are printed on the y-axis label of smoothing plots, as shown in Figures 39 and 40 for timing error data, and Figures 60 and 61 for spatial error date, in their respective sections. If the edf = 1, the model closely fits the data without smoothing, indicating that the data are reasonably linear. Smoothing parameters with higher values that indicate that the data are non-linear and that more extensive smoothing is required to fit a model to the data (Numerical Algorithms Group, 2002). Thus, larger estimated degrees of freedom (*edf*) indicate that the smoothing curve is non-linear. The mgcv package in *R* software calculates the optimal *edf* value along principles similar to the Akaike Information Criterion (*AIC*), which is also used to evaluate the optimal model. Details of the *AIC* will be described in detail in Section 4.3.4.

4.3.3.2. The Generalized Additive Model (GAM) Framework

In Section 4.2, I gave a brief description of the linear model, and showed how linear mixed models were further developments of regression for repeated measures data. Generalized additive models are another extension of regression for estimating the effects of the predictor variables on nonlinear data by using smoothers. Thus, the generalized additive model, i.e., GAM, is a smoothed equivalent of the linear regression model, depicted as:

$$Y_i = \alpha + f(X_i) + \varepsilon_i \quad \text{where } \varepsilon_I \sim N(0, \sigma^2) \qquad \text{(Equation 7)}$$

The model in Equation 7 contains the overall mean (α); the term $f(X_i)$, which denotes a smoother applied to the explanatory variable(s) at a particular observation (*i*); and ε_I, denoting the residuals (i.e., errors) representing deviations from the regression line determined by the model. Because of using the smoother, we cannot quantify the relationship between X_i and Y_i with a simple formula and rely on the values of the fixed effects and the smoother at a particular point to estimate values.

Generalized additive models (GAMs) can be more complex than the model depicted by Equation 7. GAMs may include multiple smoothers or sets of fixed and smoothed variables that have a linear relationship to the outcome variables. For example, Equation 8 represents a hybrid equation for a semiparametric GAM in which the predictors include one smoothed variable $f(X_i)$ and an interaction between two explanatory variables, *factor* (W_i) and *factor*(Z_i):

$$Y_i = \alpha + f(X_i) + factor(W_i) * factor(Z_i) + \varepsilon_I \qquad \text{where } \varepsilon_i \sim N(0, \sigma^2)$$

(Equation 8)

Note that the model depicted by Equation 8 contains fixed effects only, and does not include a random component. The generalized additive model (GAM) as shown in Equation 7 and Equation 8 is limited to data containing independent observations. Thus, the GAM has the same limitations as the linear regression model when data are correlated: it ignores the dependence structure in the data. An extension of the GAM, the generalized additive mixed model (GAMM) is needed to account for nonlinearity and correlated data structures.

4.3.3.3. The Generalized Additive Mixed Model (GAMM) Framework

The generalized additive mixed modeling (GAMM) procedure effectively handles nonlinearity in correlated data structures by combining the GAM approach, which applies smoothers to deal with nonlinearity, with a linear mixed model framework for correlated data. Once the researcher has determined that a mixed model is necessary due to dependence in the data, the researcher must select fixed and random predictor variables. The result is a model containing a fixed part and a random part, as described earlier by Equation 1.

$$Y_i = \text{fixed variable(s)} + \text{random variable(s)} \qquad \text{(Equation 1)}$$

As is true with linear mixed models, a variable such as *participants* is a random factor when participants produce multiple observations and when we intend to generalize the results to the population from which the participants were selected.

A mixed effects model requiring smoothers and including nested random effects is a special case of the generalized additive mixed model. Recalling the earlier discussion of mixed effect models with nested random effects, the notation for this type of model is described by Equation 4.

$$Y_{ijk} = \alpha + a_i + b_{j|i}X + \varepsilon_{ijk} \qquad \text{(Equation 4)}$$

Equation 4 can be applied to a generalized additive mixed model containing the overall mean, α, and a random intercept a_i for each participant$_i$, and a random intercept $b_{j|I}$ for each trial associated with each the corresponding participant (in statistical terminology, trial$_j$ nested in *participant$_i$*) assuming equal effects for each predictor (i.e., equal slopes) and the addition of smoother(s). Applied to the present data, this model depicts variability by three separate indicators: (a) σ_a^2 denotes the variability between participants; (b) σ_b^2 denotes the variability between trials; and (c) σ_ε^2 denotes differences between the errors at individual points within each trial and the mean errors at those points. Note that the specific explanatory factors are not included in this skeletal structure.

Following a determination of fixed and random factors, the researcher must determine whether a smoothed model is appropriate, and, in particular, which fixed predictor variables require smoothing. Choosing variables to smooth is determined by assessing exploratory plots of each predictor variable to discover whether nonlinearity is present. After viewing the exploratory plots, some experimentation applying smoothers to one or more explanatory variables is needed to evaluate their performance with and without the smoothers. The resulting model combines features of mixed models, which account for dependence in the data, with a nonlinear functional form that permits analysis of nonlinearity when evaluating relationships between predictor and outcome variables (Keele, 2008).

In the present data set, smoothers were not needed to evaluate wiggle in the writing line, but they were necessary for proper analysis of timing and spatial errors.

In the timing error data, the full models included interactions between fixed effects Feedback (two levels: absent and present), Target Rate (two levels: changing and constant), and Gender (two levels: female and male) with a smoother applied to the data to interpolate between peaks with a continuous smooth line, eliminating roughness in the line plotted across the samples. Random effects included the unique trial, nested in participant, as shown in Equation 4. In the spatial error data, the full models included interactions between fixed effects Feedback (two levels: absent and present), Target Rate (two levels: changing and constant), and Gender (two levels: female and male), with a smoother applied to each peak-to-peak segment in the samples. Random effects included the unique trial, nested in participant. Thus, the timing error data and spatial error data were analyzed with the same skeletal model structure shown in Equation 9. Smoothed variables are indicated by the italic letter *f*; the subscript in f_{ftg} *(peak)* stands for the smoother applied to peak which is allowed to have different values based upon the Feedback condition (*f*), Target Rate condition (*t*), and Gender. Thus, equation 9 includes the response from the fixed variables of Feedback, Target Rate, Gender, and the smoothed term at each Peak; random variables are denoted by a_i and $trials_{j|I}$, which allow for intersubject differences in repeated responses from participants, with trials nested in participants.

$$Y_{ijk} = \text{Feedback}_{ijk} \ * \ \text{TargetRate}_{ijk} \ * \ \text{Gender}_{ijk} + f_{ftg}(Peak) + a_i + trials_{j|i} + \varepsilon_{ijk}$$

(Equation 9)

The index f_{ftg} *(peak)* stands for the smoother applied to peak, allowing it to have different values based upon the Feedback condition (*f*), Target Rate condition (*t*), and Gender.

Since statisticians use a wide range of indexing systems to describe the hierarchical structure of data in generalized additive mixed effect models, I will outline the details of the system used in Equation 9 as follows:

(1) The indices *ijk* denote the *k*th observation from the *i*th participant ($i_{1\text{-}35}$) with the *j*th trial ($j_{1\text{-}16}$) nested in the *i*th participant. The *k*th observation refers to the error obtained at one of the peaks or peak-to-peak segments in the sample.

(2) The index f_{ftg} *(peak)* stands for the smoother applied to peak, allowing it to have different values based upon the Feedback condition (*f*), Target Rate condition (*t*), and Gender (g). Smoothers, indicated by f_{ftg} before (*peak)*, are not a parametric term in a model, but a line in a graph that represents an optimized participant response at each successive peak.

(3) The random effects are designated by the terms a_i and $trials_{j|I}$ which allow for separate intercepts for participants (a_i) and trials nested in participants (trials$_{j|I}$);

(4) The errors, ε_{ijk}, are obtained from the differences between the *k*th observation and the corresponding fitted values for the observations in the *j*th trial, nested in the corresponding *i*th participant.

Please note that the same structure would apply to the random slope model, if different coefficients were used to designate slopes. An additional term would be needed if both random intercepts and random slopes were included in the model.

The skeletal structure in Equation 9 does not provide any details on variance stabilizing functions that may be selected to allow for separate error variances. A close examination of model syntax is needed to determine whether covariance structures were manipulated. Changes in correlation structures may be chosen by the researcher when indicated by features of the data, such as non-constant variance or special correlations among subsets of the data, or for theoretical reasons. However, these manipulations are not indicated by the skeletal structure of the model shown in Equation 9. After these steps have been completed, the researcher should test the model again for heteroscedasticity. If the residuals are reasonably random, we can assume that the model accounts for nonlinearity in the data, and we can proceed with model testing and simplifications. In many types of data, however, nonconstant variance is typical. With classic ANOVA models, nonconstant variance may be handled by a variety of procedures, including Box-Cox transformations and adding exponential terms to the model. Transformations are not defined for mixed effects models (Zuur, personal communication, October 14, 2008), but fortunately several other procedures can be used to induce constant variance, or accommodate particular types of non-constant variance. For example, mixed effect models allow for using different variance components for different variables or other covariance structures than compound symmetry, which is most common. In the present study, a procedure denoted in *R* as *varIdent* was used, allowing each factor level of specified variables to have different variances.

As in linear mixed effect models, intraclass correlations (ICC) should also be applied in GAMMs to assess the contribution of variance from different subgroupings in the model. Optimal model structures can be assessed with the Akaike Information Criterion (*AIC*) for linear mixed models and GAMMs, which will be described in detail in Section 4.3.4.

4.3.4. Determining Optimal Model Structures

As described in Sections 4.3.2 and 4.3.3, there are many mixed effect models from which to choose when modeling repeated measures data. Thus, it is important to have a standard for selecting an optimal model from possible alternatives. Dayton (2003) advocates the use of the *Akaike Information Criterion* (Akaike, 1974) for choosing optimal models among competing alternatives. The *AIC* is a numerical measure of model fit and model complexity, where model fit is typically expressed as a function of the log likelihood and model complexity is defined in terms of parameters (Akaike, 1974; Zuur et. al, 2007). The *AIC* is calculated by

$$\text{AIC} = -2*\log(\text{Lik}) + 2n_{\text{par}} \qquad \text{(Formula 7)}$$

where log(Lik) represents the log-likelihood of the model and n_{par} represents the number of parameters in the model.

The output from *R* reports both the *AIC* and *BIC,* the Bayesian Information Criterion (Schwarz, 1978). According to Littell, Milliken, Stroup, Wolfinger, and Schabenberger (2006), the *AIC* chooses more complex models than the *BIC* in repeated measures analyses and is preferred when Type I error control is more important than retaining statistical power. Dayton (2003) advises researchers to select the model that achieves the lowest *AIC* value from alternative models, stating that "…among the models under consideration, it can be argued that the preferred model (i.e., min(AIC) model) has the smallest expected loss of precision relative to the true, but unknown, model" (p. 284). In the same article, Dayton claimed that information models such as the *AIC* are superior to *F*-test techniques which involve an accept/reject decision between two alternate models. Further, Dayton proposed that the *AIC* is more accurate than *F*-tests techniques and other information models such as the Bayes Information Criterion (*BIC*) or the Consistent Akaike Information Criterion (*CAIC*), particularly for complex models (p. 285), although that view is not uniformly held (Oehlert, personal communication, June 11, 2008).

Following the guidelines set forth by Littell et al., (2006) and Dayton (2003), I used the Akaike Information Criterion (*AIC*) to select the most parsimonious and best fitting models that represented the data in the present study. Model comparisons with the *AIC* were first used to determine whether the mixed effects models were superior to the fixed effects models. Then I used the *AIC* for model simplification where non-significant terms were removed from the *beyond-optimal* or saturated model one at a time in a sequential process, and the *AIC* values of each model were compared to the previous fuller model (Crawley, 2007; Zuur, 2007) until the *AIC* no longer decreased.

In the following sections, I will describe the analysis for each outcome variable assessed in the present study (i.e., the percentage of wiggle in the writing line, timing error, and spatial error), which are arranged in this chapter in order of complexity.

4.4. Exploration of Wiggle in the Writing Line

4.4.1. Characteristics of Wiggle Data

In this study, the definition of "wiggle" is the percentage of the trace that moved in alternating clockwise and counterclockwise directions over a sample, resulting in an undulation that appeared as a slight "zigzag" or "wrinkle" in the recorded trace. Further details can be found in Chapter 3, *Method of calculating undulations in the writing line (wiggle).* The wiggle in each sample was defined as the percentage of the sample that contained decreasing angles where increasing angles were expected. A wiggle would be produced if the tangent angle increased in a clockwise direction followed by a corrective counterclockwise direction. Each sample drawn by a participant had a unique number of sampling units, and therefore, a unique number of tangents. A percentage of clockwise motions in a sample was used to estimate wiggle since each trial produced a different number of sampling points.

In the present study, each participant contributed 16 values denoting percentage of wiggle produced in four trials performed under four separate experimental conditions (4 samples x 4 experimental conditions) as depicted in Figure 17.

Wiggle data for each participant were correlated at two levels:

(1) The four wiggle scores produced by a participant within each specific experimental condition would be expected to be more highly correlated than scores across all experimental conditions.

(2) Wiggle scores from the entire set of sixteen trials produced by each participant were expected to be more highly correlated with one another than scores across all participants.

These relationships will be depicted graphically in Section 4.4.3.

4.4.2. The Starting Point: Exploratory Data Analysis with Graphical Displays

As described in Section 4.1.3.3, graphical displays in exploratory plots are the first step in an analysis of the data. Exploratory plots facilitate efficient and logical statistical decision making (Zuur et al., 2007) by visualizing the structure of the data and the relationships between the predictor and outcome variables.

4.4.3. Data Structure Diagram

To begin the discussion of the fixed and random part of a random effects model, I will begin by presenting the hierarchical structure of a two-level mixed effects model as shown in Figure 17 with trials (observations) nested in participants. Then I will present exploratory plots for the wiggle response. These plots will prove useful in developing the statistical model for analysis.

The hierarchical two-level model shown by the data diagram in Figure 17 has two components reflecting two sources of variation for each participant: the first source of variation arises from differences between observations taken from trials performed within each experimental condition; the second source reflects differences among individual participants. To preserve clarity in Figure 17, only one participant is shown, with observations nested in each participant.

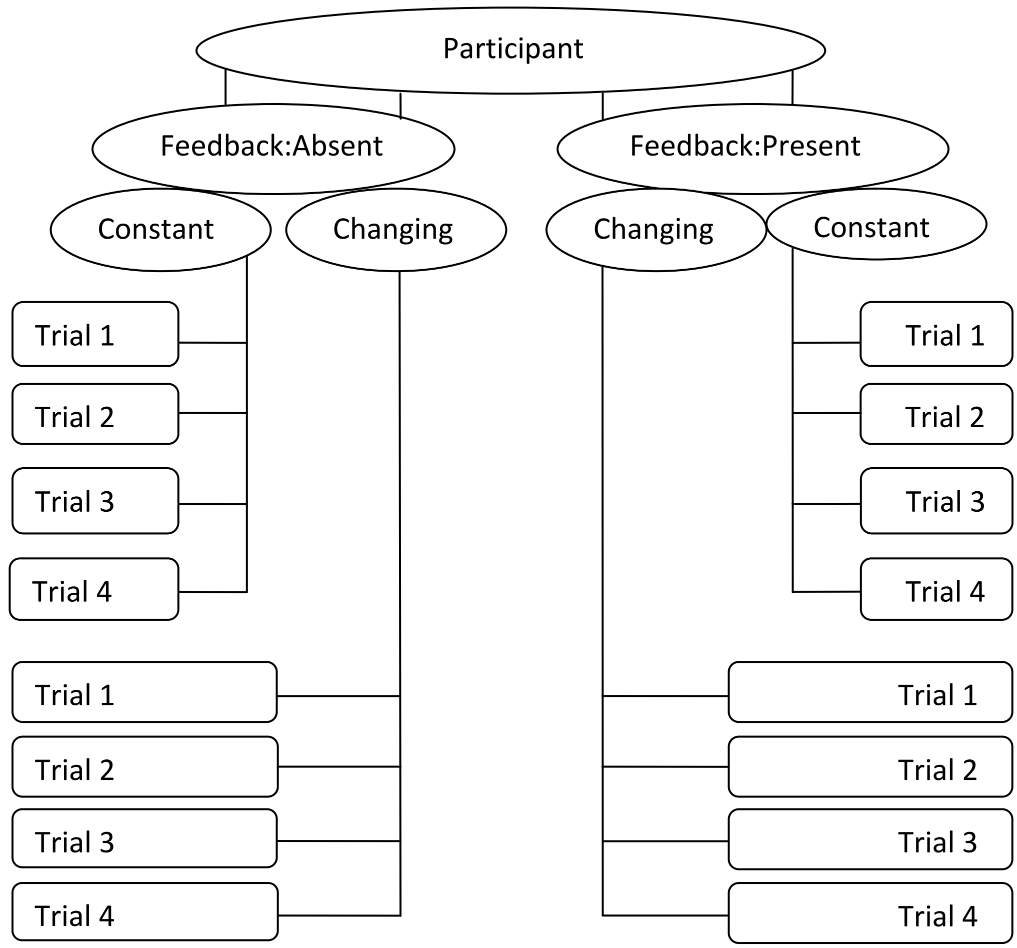

Figure 17. Arrangement of four trials from four experimental conditions nested within each participant. For clarity, the trials from one participant are denoted in this plot.

As shown in Figure 17, each participant performed four trials in each of four experimental conditions that were formed by crossing two levels of Feedback (Absent and Present) with two levels of Target Rate (Changing and Constant). Figure 17 shows the dependence of the measures within the participant, and provides clear evidence against the assumption of independence that is necessary for standard statistical tests in traditional models.

4.4.4. Exploratory Plots: Wiggle Data

In this section, I will begin by displaying the set of wiggle responses (i.e., the percentage of wiggle in the drawn samples) by different participants to determine whether participants should be modelled as random effects. Then I will display the participants' responses over trials to identify any trends. I will conclude this section on data exploration by showing the effect of individual predictors (Feedback, Target Rate, and Gender) on the observed response.

4.4.4.1. Exploratory Plots of the Participant Wiggle Responses

Figure 18 shows the values for percentages of wiggle per sample over the 16 trials performed by each of the 35 participants in the study. Each boxplot represents the range of responses from a particular participant over 16 trials. In Figure 18, participants are numbered 1-35 and indexed on the *x*-axis.

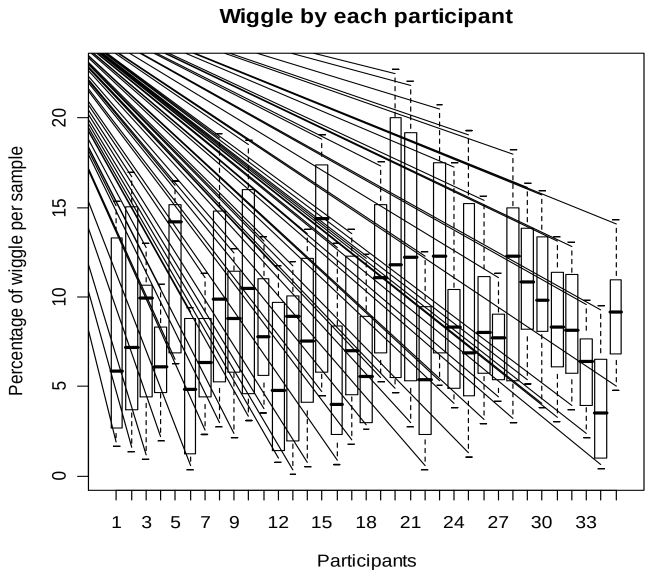

Figure 18. Boxplot of values for percentages of wiggle from 16 samples produced by each participant (*n* = 35).

The horizontal bar in the boxplot indicates the median percentage of wiggle for each numbered trial in the data set; the edges of the box enclose the middle 50% of the data; with the top edge at the 75th percentile and the lower edge at the 25th percentile of the data. The vertical lines extend to the extreme values or to 1.5 times the interquartile range, whichever is less. The single dots above or below the vertical lines represent outliers, defined as points that fall more than 1.5 times the interquartile range above the third quartile or 1.5 times the interquartile range below the first quartile (Renze, n.d.). Figure 19 extends the information displayed in Figure 18 by revealing the pattern of responses by each participant. This plot is important for determining if there are any patterns in each participant's responses that might suggest trends, such as improved responses due to motor learning or diminished responses due to constraints such as fatigue or boredom over the course of the experimental session. However, any such trends would be confounded with the four experimental conditions under which the trials were performed.

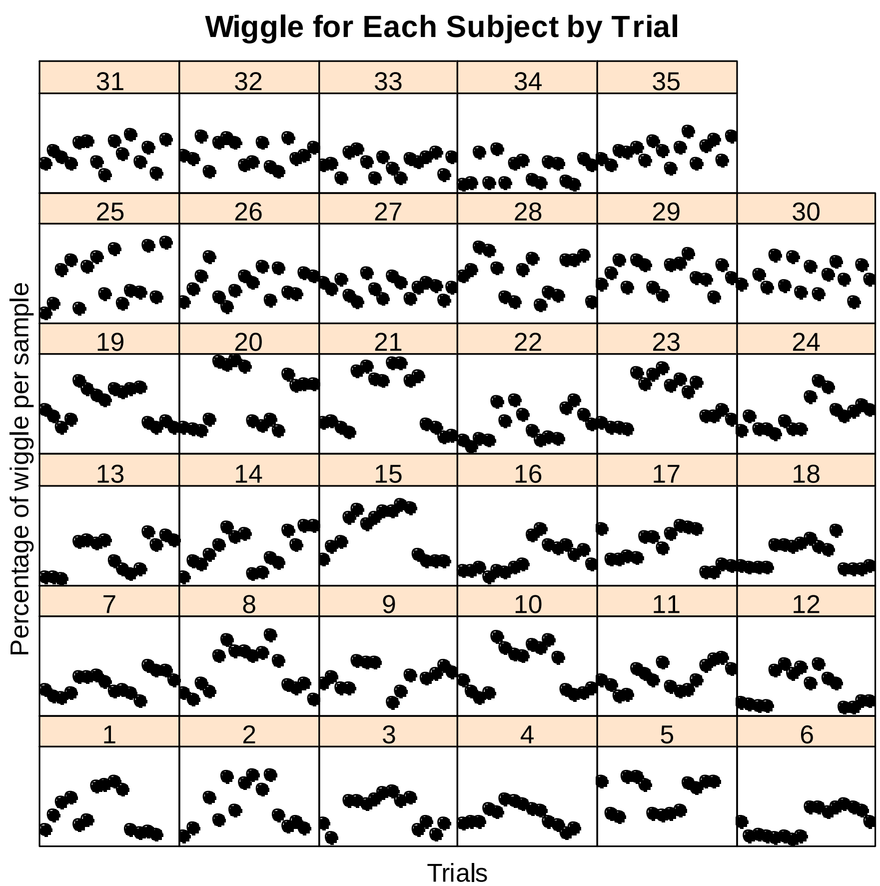

Figure 19. Lattice plot of percentage of wiggle for each trial (n = 16) produced by each participant (n = 35).

Figure 19 shows that the wiggle responses by each participant do not appear to show systematic change over time, whether by improved or diminished performance. Since Figures 17 and 18 demonstrate that the responses vary widely from participant to participant but do not reflect any systematic trend, a linear mixed model using participants as random effects may be appropriate.

4.4.4.2. Exploratory Plots of the Fixed Effect of Trial Order on Wiggle

In addition to determining whether trends appear across trials for each participant viewed separately, it is important to assess for the presence of trend across the 16 trials during the experimental session. Box-and-whisker plots can be useful for assessing systematic trends, and were used with these data to determine trends in Figure 20.

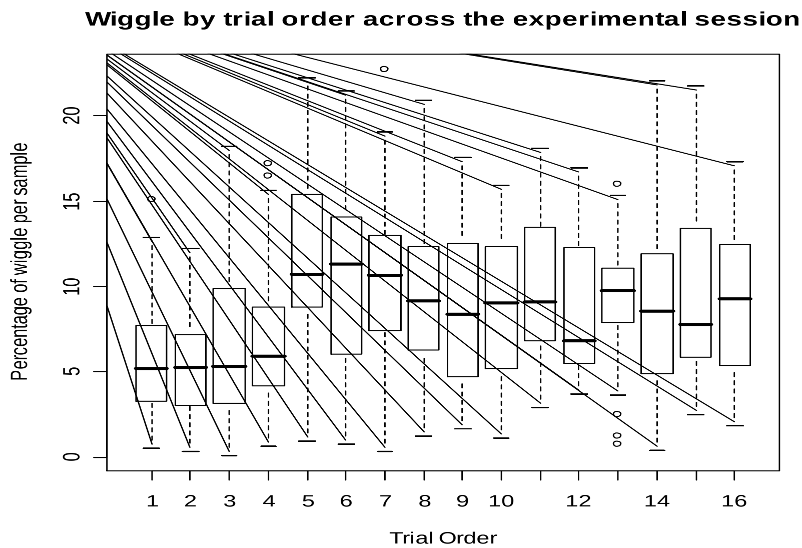

Figure 20. Mean percentages of wiggle in samples (*n* = 14-16) from all participants (*n* = 35), depicted by order of trial (1-16) across the experimental session.

Figure 20 requires some explanation since each experimental session began with two trials performed in the Feedback:Absent-TargetRate:Constant condition. The initial trials in the Feedback:Absent-TargetRate:Constant condition were used to orient participants to the task, the feel of the equipment, and the LCD display without the distraction of their own visual trace. In trials 3-16, the experimental conditions were randomized, but the randomization scheme was not completely successful, as shown by the higher number of Feedback:Absent conditions in the first trials, and the higher number of Feedback:Present conditions near the end of the series of trials.

Thus, an uneven distribution of conditions across the experimental session is reflected in Figure 20. The first two orientation trials were performed under the Feedback:Absent-TargetRate:Constant condition. A disproportionately high number of Feedback:Absent trials were performed in trial 3 and 4 as well: 63% of the scores (22 of 35 scores) came from Feedback:Absent trials in trial 3; 74% of the trials (25 of 34 trials) were Feedback:Absent trials in trial 4. Since Feedback:Absent trials were concentrated in the first four trials in the experimental sequence, a lower percentage (137 of 544 trials, or 37%) of Feedback:Absent trials were performed in trials 4-16. The only explanation for uneven distribution observed here is that the randomization process did not produce an optimal balance of trials in the two feedback conditions. Despite this situation, no trends are seen in the first four trials or across trials 5-16 shown in Figure 20, so there is no evidence that the unequal randomization fell outside the normal range for these data.

Finally, it is important to assess for the presence of trend within each experimental condition. Trends in wiggle across the entire experimental session were observed in Figure 20; in Figure 21, trends in wiggle across the trials within each experimental condition were assessed by observing whether increasing or decreasing percentages of wiggle were observed by rising or falling medians across trials 1-4 from each condition.

Figure 21 shows the percentages of wiggle for each trial within each experimental condition.

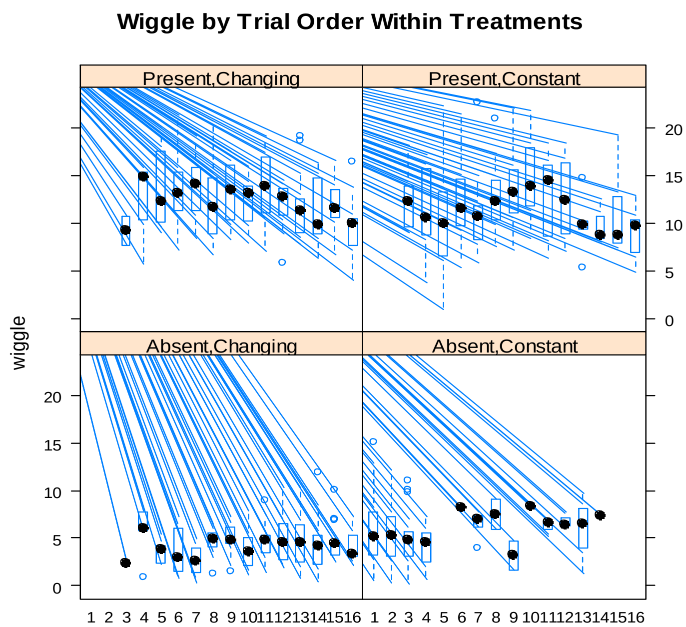

Figure 21. Boxplots of wiggle produced in each trial (*n* = 4) under each experimental condition. Boxplots represent the central 50% of the values for percentage of wiggle in each sample from trials 1-16, performed under each experimental condition. Patterns in several of these boxplots, particularly in the Feedback:Present-TargetRate:Constant condition suggest the possibility of serial correlation, meaning that the pattern of responses over time repeat themselves over a time series.

In Figure 21, the absence of a boxplot at a numbered trial, as in the Feedback:Absent-TargetRate:Constant condition indicates that the Feedback:Absent-TargetRate:Constant experimental condition was never chosen for the fifth trial during any experimental session. No apparent upward or downward trends are visible in the lattice plot of observed wiggle responses across trials in each experimental condition, suggesting that the median wiggle did not increase or decrease over the trials performed in each experimental condition.

4.4.4.3. Exploratory Plots of the Fixed Effects of the Experimental Conditions on Wiggle

In Figure 22, the percentages of wiggle in samples produced under each experimental condition are displayed. These plots are a preliminary view (from 35,000 feet, so to speak) which does not account for intersubject and other variability.

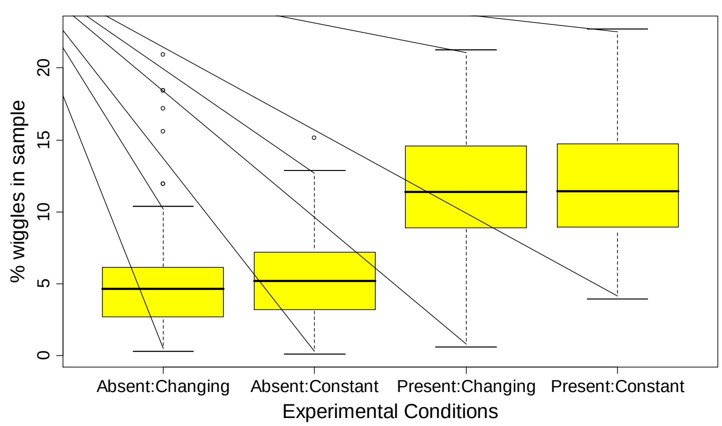

Figure 22. Percentage of wiggle in samples (*n* = 129-140) scores for each condition performed in each experimental condition. The area within the box represents the central 50% of the values for percentage of wiggle in the samples performed under each condition. The edges of the boxes indicate 25th (lower edge) and 75th (upper edge)

percentiles (i.e., the interquartile range), respectively. Vertical dashed lines extend 1.5 times the interquartile range in both directions, which is roughly two standard deviations above or below the mean (Crawley, 2007).

Figure 22 shows a higher percentage of wiggles in the Feedback:Present conditions than in the Feedback:Absent conditions. These boxplots also reveal a slight interaction of Target Rate (changing and constant) with the two Feedback conditions (Feedback:Present and Feedback:Absent). Thus, the output from the statistical analysis is expected to show a difference in the effect of the two Feedback conditions. Differences in the effects of Target Rate are less clear, and may depend on the level of Feedback.

4.4.4.4. Exploratory Plots of the Fixed Effects of Gender on Wiggle

In Figure 23, the effect of the third predictor, Gender, is displayed as a lattice plot with a division for each experimental condition.

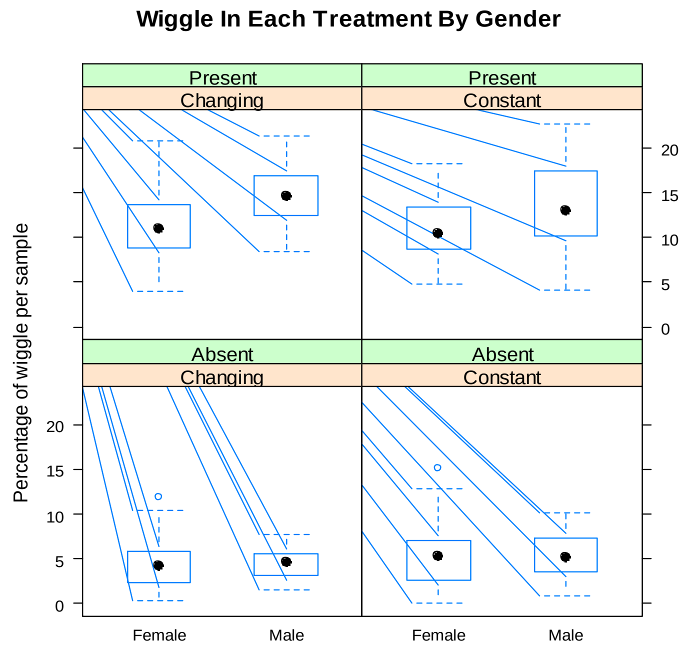

Figure 23. Percentage of wiggle in each sample by each experimental condition (Feedback and Target Rate) and Gender. Each box in the lattice represents an experimental condition and contains two boxplots, one for female, and one for male participants.

Figure 23 shows that the effect of Gender depends on the combined factors of Feedback and Target Rate that comprised each experimental condition. When Feedback is present, males produced higher percentages of wiggle than did female participants in both changing and constant Target Rate conditions. In contrast, the mean percentage of wiggle appears generally the same for males and females samples when Feedback is absent. In Feedback:Absent conditions, the range of wiggle is narrower for males than for females.

4.4.4.5. Summary of Wiggle Data from Exploratory Plots

The exploratory plots in Figures 19-20 indicate that individual participants performed quite differently from one another, suggesting that a linear mixed model with participants as a random effect is appropriate. Variation in the percentage of wiggle across trials seemed to vary from participant to participant (i.e., some participants had a narrow range of variability, while others had much broader ranges). These observations provide additional support for the decision to model participants as random effects in the initial full (i.e., *beyond-optimal*) model for wiggle. Wiggle responses differed for each level of Feedback, Target Rate, and Gender, indicating that the *beyond-optimal* statistical model for these data must include each of these factors and the interactions among them as fixed effects.

Figures 19, 20, and 21 showed no apparent trends in the percentage of wiggle in samples performed over the course of the experiment. Although it may be tempting to dismiss the role of practice in repeated trials based on these plots, statistical analysis and review of residuals from the statistical models will be necessary to make a determination about the effects of repeated trials on the outcome.

The patterns of the data shown in Figures 22 and 23 suggest a possible interaction among Feedback, Target Rate, and Gender. This result indicates that the initial model for the percentage of wiggle in the handwriting samples must contain a three-way interaction among these factors, even though the interaction between

Feedback and Target Rate shown in Figure 22 appears to be slight. Thus, the initial linear mixed effects model will include fixed effects formed by a three-way interaction between Feedback, Target Rate, and Gender; random effects will include participants, with trials (in sequence, from 1-16) nested in participants.

4.5. Application of the Linear Model to Wiggle in the Writing Line

The discussion in Sections 4.2 and 4.3 provided a foundation for the procedures that may be used in the statistical analysis of repeated measures data. The model building process requires an assessment of the goals of the experiment and structure of the data. After examining the exploratory plots, the researcher develops appropriate models, using terms that reflect the variables of interest, the design of the experiment, and the information sought from the data (Keele, 2008).

For the analysis of wiggle, I was interested in learning about the effect of concurrent visual feedback on the incidence of wiggle, the stability of the feedback effect when a timing perturbation occurred, and the effect of gender. The model was developed after reviewing graphical displays of the data, Figures 22 - 23, which showed interactions among Feedback, Target Rate, and Gender, and Figure 17, which showed the structure of the data.

4.5.1. Step 1: Evaluating the Need for a Linear Mixed Model

West et al. (2007) and Zuur et al. (2009), describe a protocol for model building in the linear mixed effect framework. These authors advise the researcher to perform a statistical comparison of fixed and mixed effect models to determine statistically whether the analysis requires random effects, or if a fixed effect model is adequate to explore the research questions. To address this question, I used a likelihood ratio test to compare an ordinary linear regression model, designated by Equation 10, to the linear mixed effects model with participants designated as random effects in a random intercept framework (a_i), designated by Equation 11, depicted previously by Figure 17.

Wiggle-gls: $Percentage\ of\ wiggle_i = Feedback_i * TargetRate_i * Gender_i + \varepsilon_i$

(Equation 10)

Wiggle-lme: $Percentage\ of\ wiggle_{ij} = a_i + Feedback_{ij} * TargetRate_{ij} * Gender_{ij} + \varepsilon_{ij}$

(Equation 11)

Thus, the linear regression model (designated by *gls* in Table 1) is compared to the linear mixed effects model (designated by *lme* in Table 1).

Table 1

Comparison of Wiggle Data Using Generalized Least Squares and Linear Mixed Effects Models

	Model	df	AIC [a]	BIC[b]	logLik	Test	L.Ratio	p-value
Wiggle-gls	1	9	2805.368	2843.925	- 1390.684			
Wiggle-lme	2	10	2569.426	2612.267	- 1274.713	1 vs 2	237.9421	<.0001

The output from *R* is reported with information on the *AIC* and *BIC,* which assess model fit. The term *logLik* in the output represents the value of the logarithm of the likelihood, or the product of the values of the probability density function for a given set of data. This value is used in likelihood ratio (L.ratio) tests to compare the fit of two models to the data. *Test* refers to the models that are being compared. In Table 1, *Test* indicates a comparison between Model 1 and Model 2. *L. Ratio* refers to the Likelihood Ratio test, where the value of the likelihood ratio is compared to a chi-square distribution with degrees of freedom equal to the difference of degrees of freedom between the more complex and the simpler model (Model 2 df – Model 1 df, in this case). The likelihood ratio is compared to the chi-squared distribution for the appropriate degrees of freedom (Model 2 df – Model 1 df) to determine whether the difference between the two models is statistically significant. The result of the likelihood ratio test is reported as a *p-value,* which indicates the probability that the models being compared would produce likelihood ratios as extreme as these, if the null hypothesis (i.e., no statistically significant difference between models) is actually true. If the differences between the models are statistically significant, the model with the best fit is chosen on the basis of the AIC or BIC values, where lower values indicate a better fit.

Table 1 shows that the linear mixed effects model designated by wiggle-lme is significantly different ($p < .0001$) from the fixed effects model designated by wiggle-gls. The lower AIC value for the linear mixed effects model compared to the general linear fixed effects model (2569.426 vs. 2805.368) suggests that the linear mixed effects model is superior to the standard fixed effects model for providing a more complete and accurate summary of the data. Thus, the use of the linear mixed model to analyze these data appears to be justified.

4.5.2. Step 2: Initial Modeling Procedures: Developing a Beyond-Optimal Model

After determining that the linear mixed model was appropriate, I developed the first model based on the research questions about the effect of visual Feedback on wiggle and the stability of that effect, the data structure, and the exploratory plots as discussed in Section 4.3. Eleven data points were removed as outliers after reviewing the graphical displays of the data in Figures 18 – 23. Outliers were defined as points that fall more than 1.5 times the interquartile range above the third quartile or 1.5 times the interquartile range below the first quartile (Renze, n.d.).

I applied the skeletal framework of the linear mixed model, i.e.,

Y_i = fixed variable(s) + random variable(s) (Equation 1)

to the wiggle data in the present study, with terms representing the explanatory factors (Feedback, Target Rate, Gender, and Trial within conditions) and the hierarchical structure of the data with participants modeled as random effects, with trials nested in participants. Therefore, the beyond-optimal model contained these fixed factors:

(a) Feedback information with two levels (absent and present);

(b) Target Rate with two levels (constant and changing);

(c) Gender with two levels (female and male)

The beyond-optimal model also included participants as a random effect.

The model allowed for the possibility that the effect of any single explanatory variable, such as Feedback, may depend on a second covariate, such as Target Rate or Gender, producing an interaction across predictors. For example, Figure 21 shows that the percentage of wiggle in the tracings changes with different levels of Feedback and Target Rate. Figure 23 shows that the percentage of wiggle in each sample in the Feedback:Present condition has different values for different genders. Thus, the fixed portion of the model (in words, after Galwey, 2006) is represented as:

Feedback * Target Rate * Gender + Trial $\text{Number}_{\text{conditions}}$ + Trial $\text{Number}_{\text{session}}$

(Formula 8)

The symbol * signifies an interaction among three potential predictor variables which may change at different levels. This model structure is a bit of shorthand for a fully detailed model. Since terms enter the model hierarchically, all two-way and main terms are included in the model when a three-way interaction is present according to a mathematical rule. Trial Number is evaluated separately to assess changes over the repetitions in the experimental session, and is not expected to interact with any of the other predictor variables.

Combining the fixed and random part of the model and applying it to the factors in this study produces the following random-intercept model:

$$\textit{Percentage of wiggle}_{ij} = \textit{Feedback}_{ij} * \textit{TargetRate}_{ij} * \textit{Gender}_{ij} + \textit{Trial Number}_{ab} + \varepsilon_{ij}$$

(Equation 12)

with participants entered as random effects, allowing for participant-specific intercepts where *percentage of wiggle*$_{ij}$ refers to each observation (i.e., percentage of wiggle in a trial). Feedback$_{ij}$, Target Rate$_{ij,}$ and Gender$_{ij}$ refer to the overall effects of the predictors (i.e., two levels of Feedback (absent and present) crossed with two levels of Target Rate (changing and constant)) on the percentage of wiggle in each trial, for each participant, allowing for separate effects for males and females (Gender$_{ij}$). Trial order$_{ab}$ is evaluated separately to determine if trends in performance occurred over the trials in each experimental condition, $\boldsymbol{a}$ (1-4) or over the experimental session, $\boldsymbol{b}$ (1-16). Thus, the analysis will determine the percentage of wiggle in each trial, and evaluate the interaction of Feedback∗Target Rate∗Gender on the percentage of wiggle observed and assess for trend (Trial order$_{ab}$). The term ε_{ij} represents the random component of the model and refers to the errors for each trial and participant observation.

4.5.2.1. Initial Modeling Procedures: Checking For Order Effects

The effects of trial order within each experimental condition and over the experimental session were included in the preliminary models as fixed effects to determine whether the data depended on sequence. The full model with participants as the sole random effect was analyzed to determine whether trial orders within experimental condition or over the experimental session were significant predictors of observed wiggle (Equation 12).

Trial order (1-4) within experimental condition was not significant statistically (F= .98, p = .400). Trial order (1-16) over the experimental session was marginally insignificant (F= 1.63, p = 0.061). The p-values are large enough that neither trial order within each condition or trial sequence over the experimental session should be included as fixed effects in subsequent wiggle models. Since trial order and trial sequence were not found to be significant statistically, they will not be included in future discussions of the model for wiggle.

Since trial order was found to be non-significant, the *beyond-optimal* model was represented by fixed terms for the three-way interaction between Feedback, Target Rate, and Gender with Participants entered as random effects a_i allowing for participant-specific intercepts, shown in Equation 13:

$$Percentage\ of\ wiggle_{ij} = Feedback_{ij} * TargetRate_{ij} * Gender_{ij} + a_i + \varepsilon_{ij} \qquad \text{(Equation 13)}$$

4.5.3. Step 3: Initial Modeling Procedures: Checking For Constant Variance

A residual plot of the model in Equation 13 showed non-constant variance, or heteroscedasticity, as displayed in Figure 24. Heteroscedasticity is a concern since model residuals must show a random pattern without any underlying structure for us to have confidence that the model summarized the data accurately, and that the inferential tests are reliable.

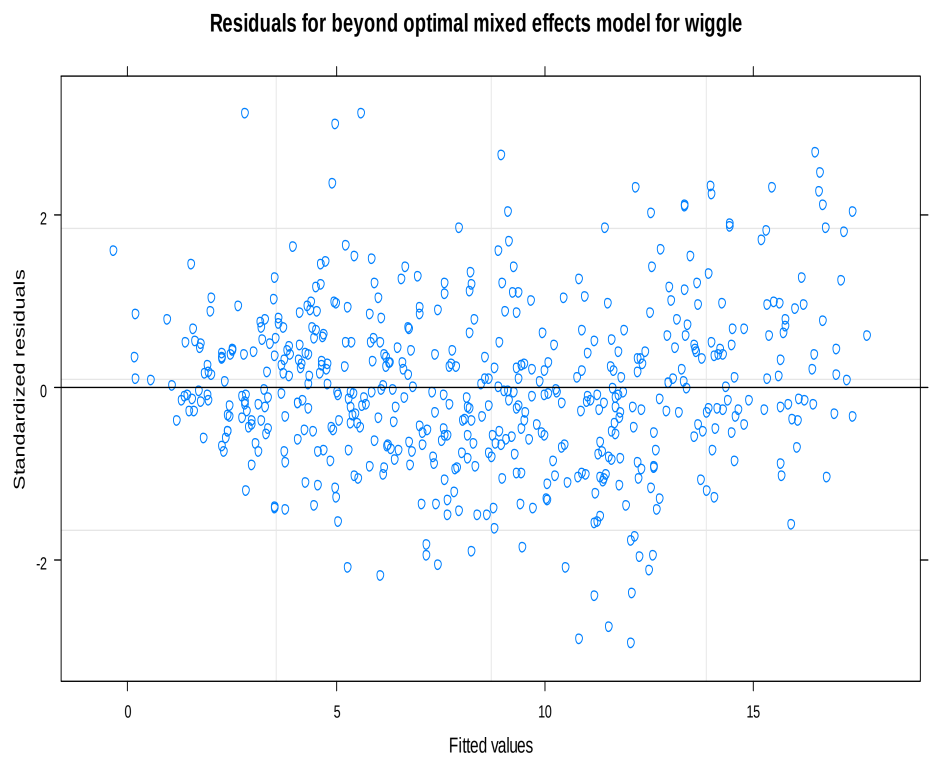

Figure 24. Standardized residuals vs. fitted values for the *beyond-optimal* linear mixed effects model for the wiggle data in Equation 13.

Standardized within-group residuals are defined by Formula 9 (Fox, 2002).

$$\text{Standardized residuals} = \frac{\hat{\varepsilon}_{ijk}}{\hat{\sigma}} \qquad \text{(Formula 9)}$$

Since standardized residuals have variance = 1, a standardized residual that is larger than two, is considered large. In Figure 24, the standardized residuals have several that exceed -2 and 2, (Draper and Smith, 1981) indicating that some outliers remain in the model. Further, the plot for this model shows not only non-constant variance, but discreteness and possible lack of fit, as seen in the upward curve on the right. As a result of these factors, methods must be selected to normalize the data so that the results of statistical tests modeling are reliable.

4.5.3.1. Dealing with Non-Constant Variance: Random Intercept and Random Slope (Random Coefficient) Models

Note that the residuals increase as the fitted values increase in Figure 24. These patterns may occur when the mixed model imposes a linear relationship on data that may actually be nonlinear or when the variance depends on the values of a predictor included in the model. The initial data model showed large differences in wiggle between conditions where the participant trace was presented concurrently with the target on the LCD monitor and when it was not. The heteroscedasticity in these data was reduced when the model allowed participant-specific effects in addition to the random intercepts (a_i) associated with each participant, as shown in Equation 14. The conventional term for a model with participant-specific intercepts and slopes is a random coefficient model.

$$Percentage\ of\ wiggle_{ij} = Feedback_{ij} * TargetRate_{ij} * Gender_{ij} + a_i + b_i + \varepsilon_{ij} \qquad \text{(Equation 14)}$$

As before, the *beyond-optimal random coefficient* model was analyzed with REML (restricted maximum likelihood).

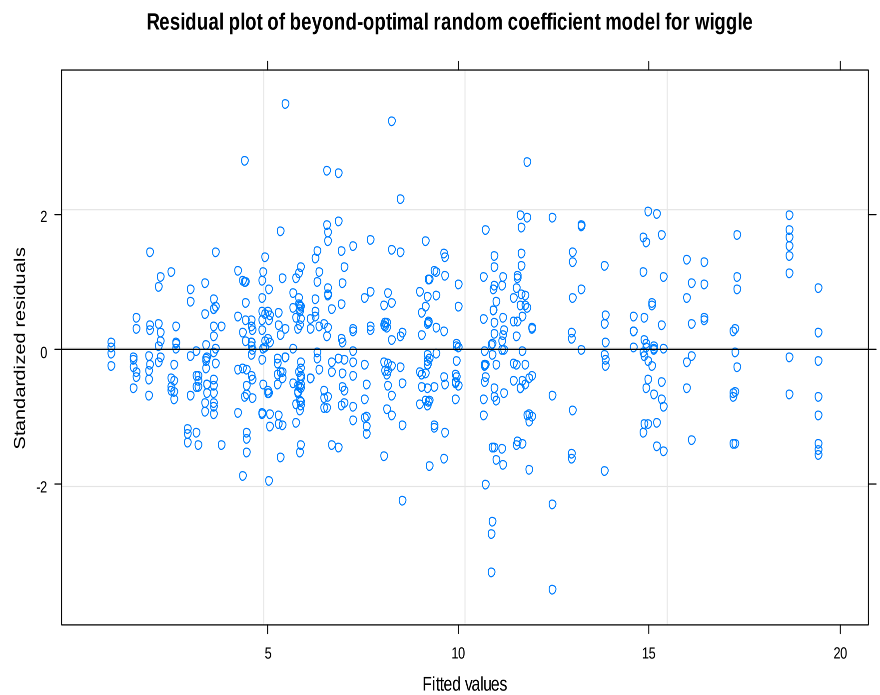

Figure 25. Standardized residuals plotted against fitted values for wiggle data, from the model specified by with Equation 14, with the addition of random intercepts for participants and a participant-specific slope for type of feedback.

Figure 25 reveals that the wiggle data residuals from the model with random participant-specific intercepts and random participant-specific slopes for feedback produced reasonably constant variance. The *AIC* values for the random coefficient model for wiggle were lower than the model containing only random intercepts for participants, indicating that the random coefficient model provided a superior fit to the wiggle data (L = 19.2779, p < .0001).

Since the random coefficient model for wiggle (Equation 14) produced randomly distributed residuals as shown in Figure 25 and a lower *AIC* than competing models, further analysis of wiggle will be based on this model. However, please note that methods for achieving constant variance in mixed models is an active area of research, and other researchers have suggested other solutions to the problem of non-constant variance in mixed effect models (Jacqmin-Gadda, Sibillot, Proust, Molina, & Thiébaut, 2007).

4.5.4. Step 4: Refining the Model for the Wiggle Data

Once a provisional model for the data is identified, the next steps involve removing non-significant predictors. After trial order was removed, non-significant terms in the model for wiggle were removed sequentially in order of highest p-value through backward stepwise elimination in the following order, refitting the model at each step:

(1) The three-way interaction of Feedback, Target Rate, and Gender was found to be non-significant ($F = 0.3005$, $p = 0.5838$).

(2) The two-way interaction between Feedback and Target Rate was non-significant ($F = 0.1237$, $p = 0.7252$).

(3) The two-way interaction between Target Rate and Gender (F = 1. 3157, p = 0. 2519) was non-significant.

(4) Target Rate was found to be nonsignificant (F = 2.0084, p = 0.1571).

After removing the terms listed in 1-4, the final reduced model contained the fixed effects of Feedback, Gender, and the interaction between Feedback and Gender, shown in Equation 15, with participants modeled as random effects with a participant-specific intercept (a_i) and participant-specific slopes (b_i):

$$Percentage\ of\ wiggle_{ij} = Feedback_{ij} + Gender_{ij} + Feedback_{ij} * Gender_{ij} + a_i + b_i + \varepsilon_{ij}, \qquad \text{(Equation 15)}$$

with participants as random effects, modeled with participant-specific intercepts and participant-specific slopes for feedback.

Since the model represented by Equation 15 can be difficult to understand in narrative form, I am including the syntax for this model in the statistical language *R,* in Formula 10:

```
Wiggle Model<- lme(Wiggle ~ Feedback*Gender,
random = list(participants = ~ Feedback), method = "REML", na.action=na.omit)
```

(Formula 10)

The syntax for the optimal linear mixed effects model for wiggle shows that the response includes the interaction between Feedback and Gender (*Feedback*Gender*). Since terms enter the model hierarchically, the interaction implies that both Feedback and Gender are included as main effects. The term *random= list(participants = ~ Feedback),* indicates that participants are entered as random effects with a participant-specific slope for Feedback (West et al., 2007). West et al. (2007) states that a random intercepts are assigned to the first term following *list* by default, indicating that the model includes participant-specific intercepts also. Thus, the syntax in Formula 10 indicates a random coefficient model for wiggle that accounts for correlations between observations from the same participant. Restricted maximum likelihood (REML) was used to fit the model.

A final graphical evaluation of the model, Figure 26, shows that the residuals for this model are homoscedastic, suggesting that this model accounts adequately for the structure in the data, although some outliers remain.

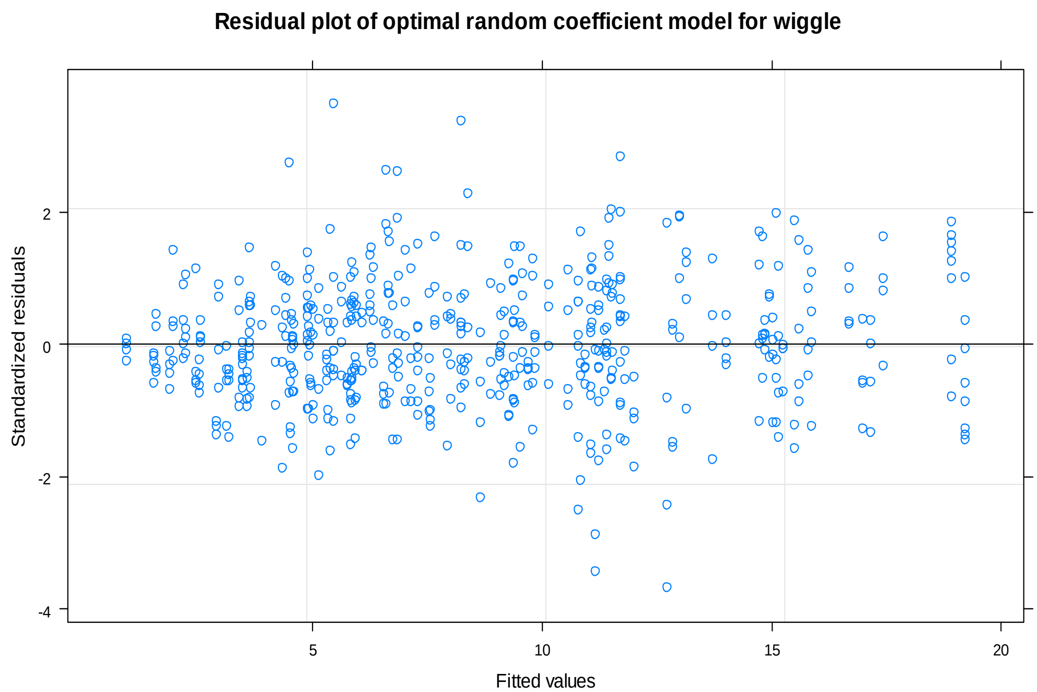

Figure 26. Residual plot for final wiggle model (Equation 15) for the wiggle data.

4.5.5. Final Tests of the Model

4.5.5.1. Akaike Information Criterion (AIC)

Optimum models are both parsimonious and efficient, containing the fewest terms needed to summarize the information in the data. In the analyses performed in this study, I chose the *AIC,* described previously in Section 4.3.4, to determine the optimum model. The reduced model with the lowest *AIC* is preferred of competing models when the p-values showed that each simplification produced a significantly different model.

In the analysis of the wiggle data, the *AIC* was evaluated as each non-significant term was removed, to ensure that the AIC decreased at each step. The pruned model with the interaction between Feedback and Gender (Feedback*Gender) and Target Rate (*AIC* = 2676.174) was compared to the model with Feedback and Gender only (AIC = 2676.219). The Likelihood Ratio test of these models indicated that these two models were not significantly different (L = 2.0449, p = 0.1527). Since the *AIC* values for the two models were different by a small amount (0.045), the p-value indicated that the difference between these models was not significant , and Target Rate was not found to be significant (p = 0.1571), I chose the reduced model with Feedback and Gender, without Target Rate, for the final (i.e., optimal) model.

4.5.5.2. Interpreting Explanatory Terms in the Presence of Interactions

When a statistically significant interaction is present, the factors in the interaction, viewed separately, do not represent a unique and meaningful main effect on the outcome (Fox, 2008), even though the statistical output of most programs provides statistical tests and p-values for those factors listed separately. Most statisticians contend that factors included in an interaction are not interpretable separately as main effects since the interpretation of one variable in an interaction depends on the level of the other factors in the interaction. In some cases, interpretations of main effects included in an interaction are possible, but limited to specific ranges or conditions in the data (Oehlert, personal communication, June 5, 2009), although these cases are rare. Underwood (1997) takes a stronger position against interpreting factors in an interaction as main effects, stating that "interpretations of main effects are impossible, or at least, unreliable when there is an interaction" (p. 321). Thus, the factors of Feedback and Gender have interdependent effects on the response that are significant statistically and therefore cannot be interpreted individually as main effects. In *Chapter 5, Results,* I show how deletion tests can be used to show that each of the predictors in this model are significant, due to the significantly poorer fit when one of them is removed.

4.5.5.3. Final Comments on the Optimum Model for the Analysis of Wiggle

In summary, the reduced model in Equation 15, with fixed effects of Feedback, Gender, and the interaction Feedback∗Gender, and random effects consisting of random intercepts for participants and a participant-specific slope for Feedback, appears to fit the data adequately. The patterns of residuals for the final optimal model shown in Figure 26 indicate that participant-specific random intercepts and participant-specific slopes for Feedback were generally effective in reducing the heteroscedasticity in the original model. Thus, these strategies appeared to be adequate choices for analyzing the wiggle data set. These results permit us to conclude that Feedback played a significant role in contributing to the amount of wiggle observed in this sample, and that the extent of the Feedback effect varied by Gender.

4.5.5.4. Intraclass Correlations

Similarities in wiggle among subgroups (i.e., within trials and Participants) were determined using intraclass correlations, $ICC_{participant}$ and ICC_{trial}, calculated using Equation 16, from the standard deviations for subjects, trials, and residuals at each peak-to-peak segment (Table 2).

Table 2

Standard Deviations for Participant, Trial, and Residuals for Wiggle Used in ICC Computations

Coefficient	Standard deviation (σ)
Participants (estimate of σ_a)	0.6595
Residuals of trials nested in participants	
(measures for each sample, estimate of σ_{ε})	0.9012

Using the standard deviations in Table 2, the intraclass correlation for scores across participants was calculated using Equation 16.

$$ICC_{participant} = \frac{\sigma^2_{participant}}{\sigma^2_{participant} + \sigma^2} = \frac{0.6595^2}{0.6595^2 + 0.9012^2} = 0.349 \qquad \text{(Equation 16)}$$

$ICC_{participant}$ was relatively low (see also the box-and-whisker plots in Figures 18 and 19), suggesting that the responses of individual participants had a broad range of responses under the four experimental conditions. This result provides additional support for the decision to model participants as random effects.

4.6. Exploration of Timing Differences between Participants and Targets

In this section, I will begin by describing the characteristics of the timing data. Following this description, I will explain the steps involved in selecting the most appropriate model, based on the discussion of the linear mixed model framework in Section 4.2 and generalized additive mixed models in Section 4.3.

4.6.1. Characteristics of the Timing Error Data

The term *timing error data* in the present study refers to measures of timing errors defined as the difference between the times that the target and participant reached corresponding landmarks on the sample. Peaks (*y(t)* maxima) were chosen as landmarks for ease of computation, as described in Chapter 3, *Segments Selected for Comparison and Analysis.* The length of time taken for the target and the participant to reach the first peak was removed to eliminate extraneous features and allow movement characteristics to be measured more precisely (Ganz, Ehrenstein, Cavonius, 1996), making the first peak the starting position (time = 0). Difference measures were obtained by computing the difference between the times that the participant drawings and the moving target stimulus reached peaks of each subsequent loop (2-12) in the template. The difference measures yielded a vector of responses for each participant sample consisting of a set of 11 values, each representing the sequential timing differences between the participant and the target at each peak, measured from the starting position at the first peak.

The research was designed so that each participant (n = 35) would contribute 176 scores to the data set, one for measured peak (n = 11) of each sample (n = 16), resulting in 6160 (35 x 176) possible values in the data set. Of the 6160 total possible scores, 196 scores (3.2%) were lost: 187 scores from 17 samples were missed due to researcher error or equipment malfunction, and nine scores were missed in nine separate samples due to participant error. As a result, 5964 timing differences (errors) between the target and participants were available for the analysis. Seven mild outliers (1.5 - 3 times the interquartile range) were identified, but they were retained in the data set as they belonged to one participant only.

The timing errors from each participant were expected to form sets of correlated observations as follows:

(1) Timing errors at each of the eleven peaks within a sample were correlated with each other.

(2) Timing errors of each participant were correlated within one another in each experimental condition.

(3) Timing errors of each participant are correlated with one another across the entire set of sixteen trials contributed by each participant.

The data structure depicting these correlations is described as a three-way nested data structure, with each subunit nested in a larger unit. In the timing data, the timing errors at each peak are nested in the sample in which they were produced. Each sample is nested within the participant who produced it.

4.6.2. The Starting Point: Exploratory Data Analysis with Graphical Displays

As discussed in Section 4.1.3.3, graphical assessments are important for facilitating logical model development by determining the structure of the data, identifying correlated observations, and exploring relationships within the data. Data exploration is particularly important when we know *a priori* that measures are correlated so that the dependent structures can be accurately understood.

4.6.3. Data Structure Diagram

In Section 4.3.2.2, *Expanding the Random Intercept Linear Mixed Model to a Three-Way Data Set*, I described a three-way data structure that can be applied to the timing error data. The timing error data requires three hierarchically arranged categories of data, including a vector of eleven responses nested in the trial in which they were performed, and a set of trials nested in the participant who performed them, as shown in Figure 27. The analysis of the timing data requires that each level of this hierarchy be reflected in the model.

The data diagram in Figure 27 displays three sources of variation for each participant:

(1) The first source of variation arises from timing errors (i.e., differences between the target and participant at the 11 corresponding landmarks) observed at each trial (one at each peak in the sample).

(2) The second source of variation reflects differences among observations taken from the four trials performed within each experimental condition.

(3) The third source reflects differences in observations within participants at corresponding peak-to-peak segments.

The data diagram in Figure 27 shows the dependence of measures by demonstrating that observations are correlated within trials, within experimental conditions (i.e., combinations of Feedback and Target Rate), and within each participant. In each timing error data set, a fourth source of variation arises from the differences among individual participants, but to preserve clarity and space in Figure 27 the data structure for only one participant is shown, and gender is not considered. Only two trials are depicted with their respective peaks, one under Feedback:Absent-TargetRate:Constant and the second under the Feedback:Present-TargetRate:Constant conditions.

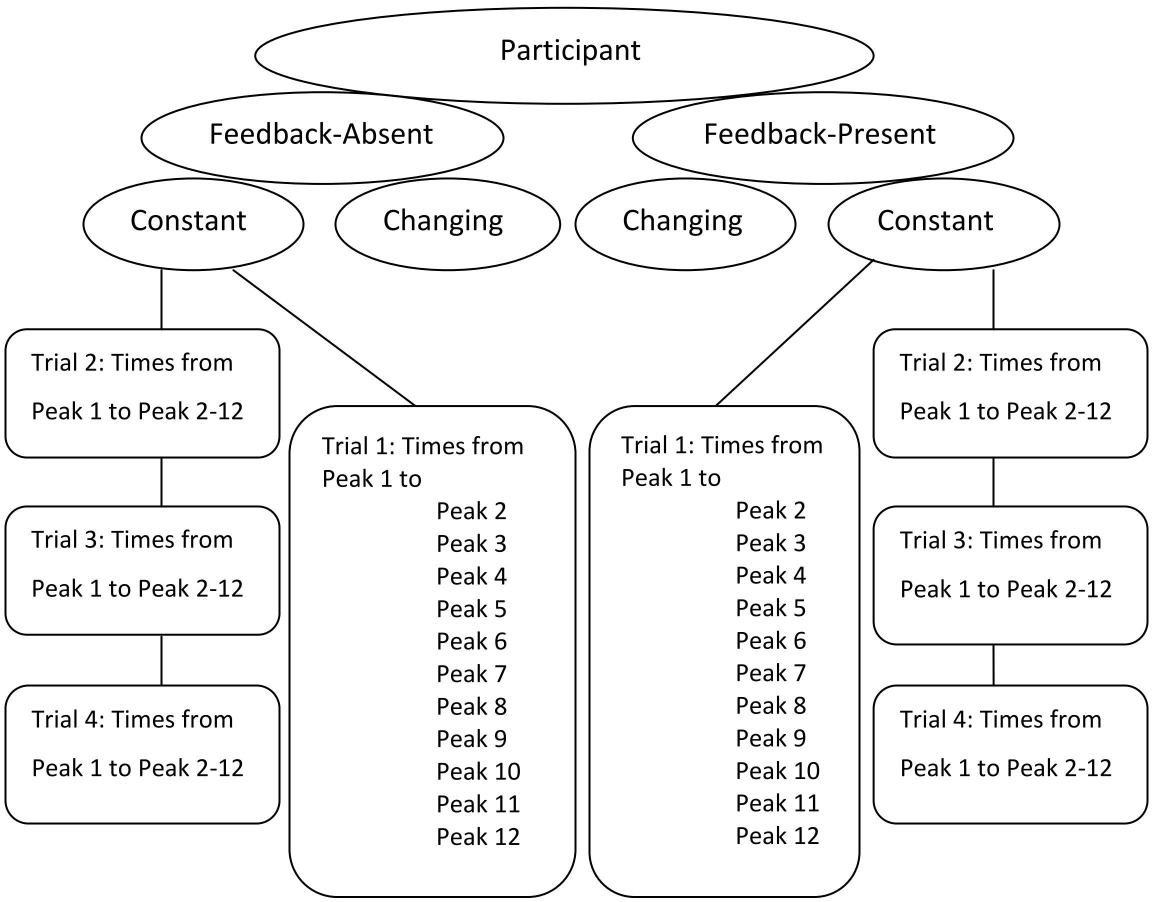

Figure 27. Arrangement of 11 peaks nested in each trial; four trials are nested in each set of four experimental conditions for each participant.

Figure 27 shows the relationship of within-subject factors by demonstrating that observations are correlated within trials, experimental conditions, and participants.

The Target Rate conditions are placed under each Feedback condition to indicate that the four experimental conditions result from crossing two levels of Feedback (absent and present) with two levels of Target Rate (changing and constant). Two trials (one Feedback:Absent-TargetRate:Constant and one Feedback:Present-TargetRate:Constant) are expanded to show the eleven peak-to-peak segments nested in each trial. The remaining TargetRate:Constant trials are denoted by boxes with the words "Trial *n*: Times from peak 1 to peak…."; trials performed in the TargetRate:Changing conditions (four with Feedback:Absent and four with Feedback:Present) are not depicted here. Gender, a between-subject factor, is not displayed here.

4.6.4. Exploratory Plots: Timing Data

In this section, I will begin by displaying the timing errors by participants, shown in Figure 28, and then display the effects of each predictor on timing errors.

4.6.4.1. Exploratory Plots of the Participant Timing Errors

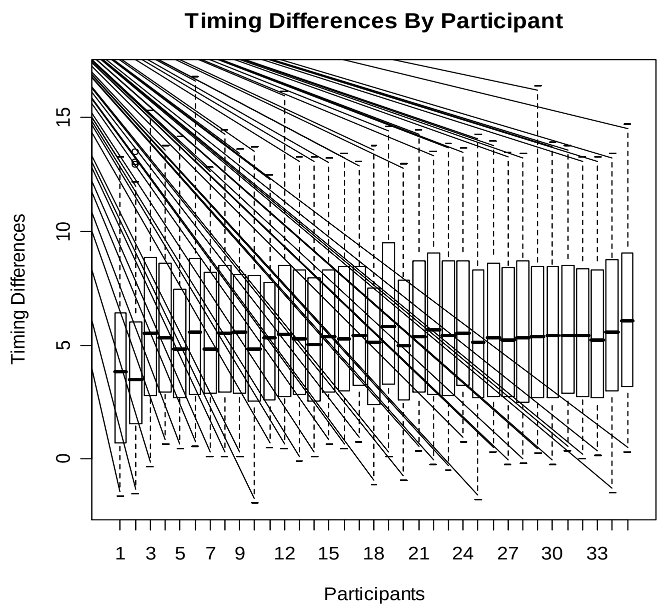

Figure 28. Pattern of timing errors for all thirty-five participants over all sixteen trials in the study.

The vertical lines in the centers of the boxplots represent the median timing error for each participant within the 16 samples each participant contributed. The medians are comparable, which may reflect the efforts of the participants to match the tempo of the moving target, or the constraining effect of the target. Despite this similarity, the variability between participants suggests that modeling participants as a random effect is appropriate.

The patterns of timing errors across the sixteen trials in the experimental session are displayed for each participant in Figure 28, with 16 box-and-whisker plots for each participant, each one representing the range of timing errors for the eleven peaks within each trial. The black points shown in the center of the boxplots represent the median timing error of a participant's 11 responses in each trial. Thus the first box-and-whisker plot under participant #1 depicts the values of the timing errors at the eleven peaks in that participant's first trial, the second box-and-whisker plot under participant #1 depicts the values of the timing errors at the eleven peaks in the participant's second trial, and so on.

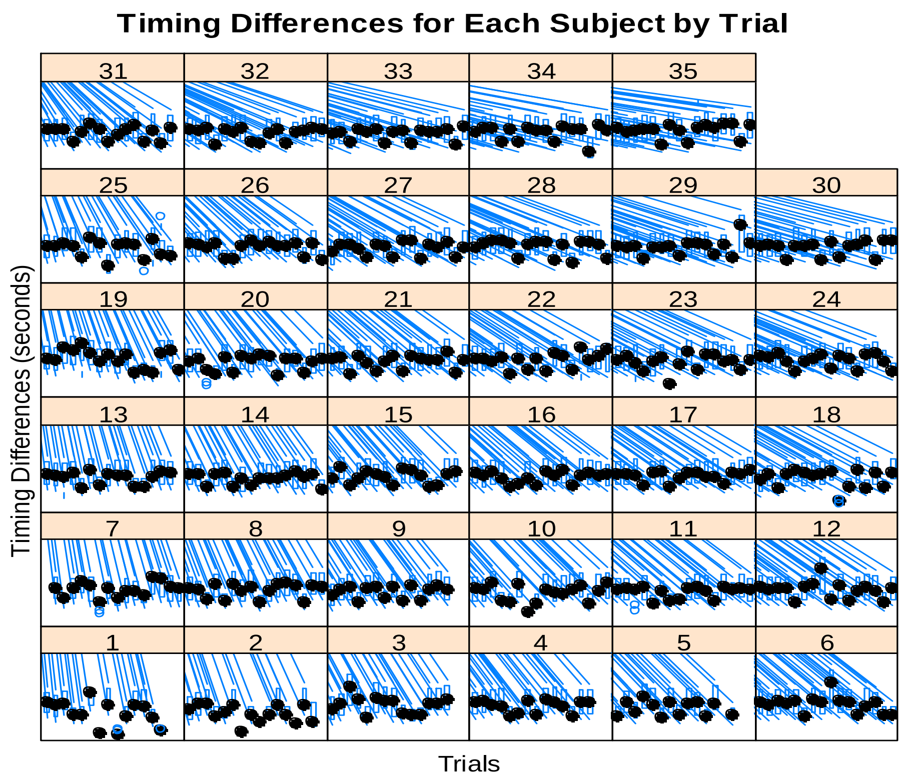

Figure 29. Timing errors for each participant for each trial. Each box (n = 35) represents a separate participant; each boxplot within the frame for a participant (n = 16) reflects the timing errors at peaks in a single trial (n = 11).

Although no trend is apparent in the box-and-whisker plots in Figure 29, it is very hard to determine whether trend is present due to the small size of the box-and-whisker plots in the lattice plot. To explore trend in a clearer display, Figure 30 represents the timing differences in 16 trials performed by a subset of six participants. These plots reveal variation across trials by the same participant, but the variation does not appear to have a systematic pattern. However, these differences are confounded with treatment effects, which must be taken into consideration when viewing these plots.

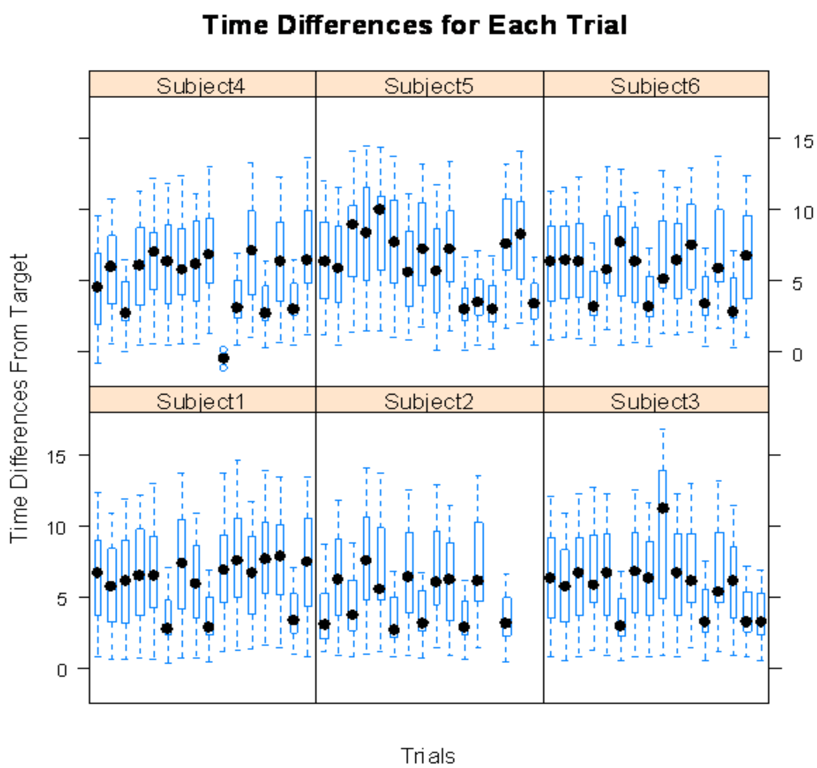

Figure 30. Timing errors by six individual participants. Each boxplot (n =16) within the frames reflects the timing errors at peaks in a single trial (n = 11) for the respective participant.

The range of responses within each trial represented by each boxplot in Figure 30 denotes all the timing differences between participants and targets at successive peaks in the trial represented by the boxplot. Note that three trials for Participant 2 (trial # 13, trial #15, and trial #16) were unusual in that those trials were missing, perhaps due to equipment malfunction. Despite these missing trials, the variability in Figure 29 and Figure 30 reveals no evidence of systematic change from trial 1 to trial 16 over the experimental session. Thus no additional measures, such as detrending, appear to be needed in this analysis.

The exploratory plots of responses from individual participants (Figures 28, 29, and 30) revealed differences in medians for the individual trials performed by each participant, and justify the decision to test participants as random effects in the model. Since trials within participants vary as well, as seen in Figure 29 and 30, the model for timing data should include trials nested in the corresponding participant as random effects in the model as well.

4.6.4.2. Exploratory Plots of the Effect of the Fixed Effects of Trial Order

In Figures 29 and 30, the responses of each trial were displayed, with a separate plot for each individual. In Figure 31, boxplots representing the responses at each trial are shown for all participants. Presumably, if responses improved or degraded over time for all the participants, the responses in the initial trials would be greater or less than the final trials, showing a trend across the 16 trials for the entire group of participants.

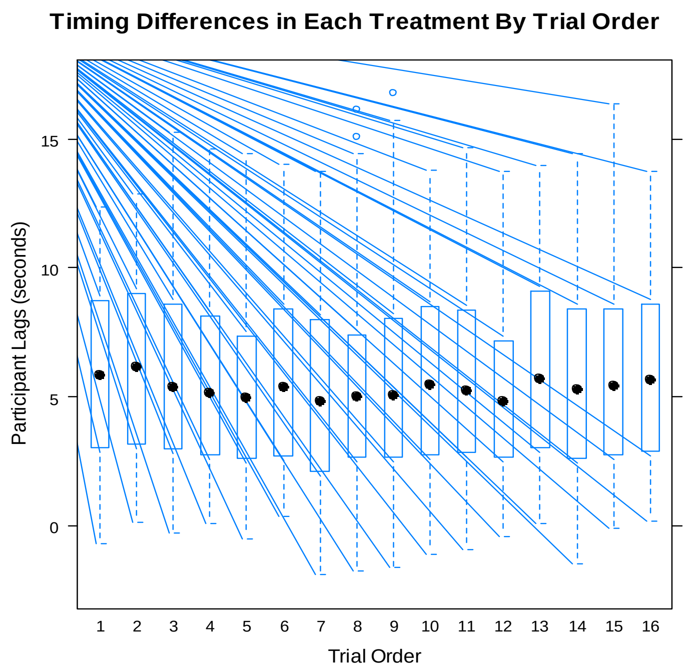

Figure 31. Timing errors in each trial shown by trial order.

Sixteen trials were performed by each participant under experimental conditions presented in randomized order. The black dots represent the median timing error of the trials performed in the order indexed on the *x* – axis (*n* = 16) by each of the participants (*n* = 35). Although Figure 31 reveals no systematic changes in timing errors across the randomized trials in the experimental session, it is important to recognize that each plot includes variation from many sources, including individual differences and differences in experimental conditions, which may mask the effect of trend. In Figure 32, I present differences in response for each experimental condition, but the same caution about multiple sources of variation in each condition applied to these plots as well.

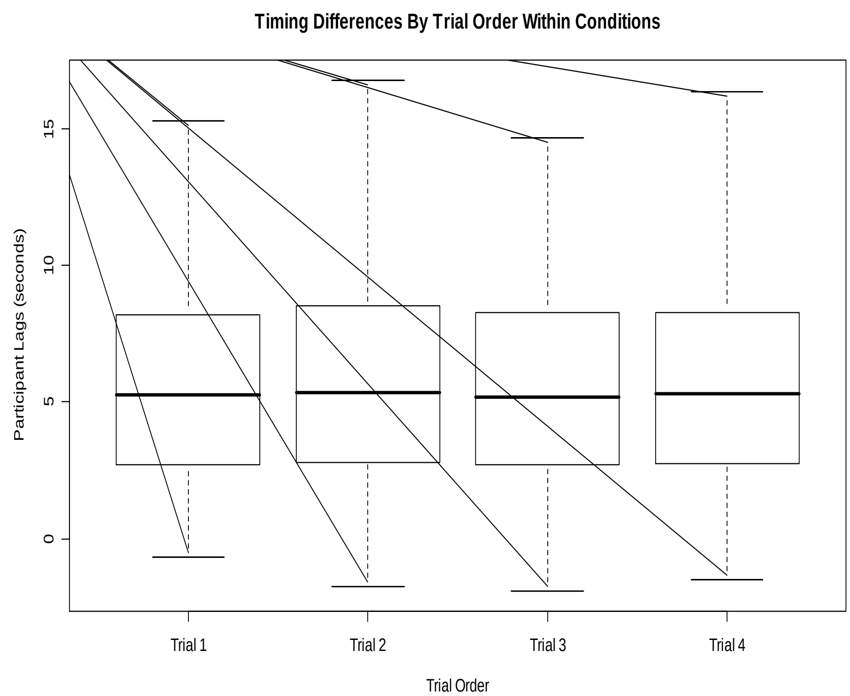

Figure 32. Timing differences between participants and targets (i.e., timing errors) by trial order (1-4), i.e., the first (n = 1527), second (n = 1536), third (n = 1528), and fourth (n = 1383) trials of each experimental condition.

Figure 32 depicts timing errors by trial order (trial 1-4) within each experimental condition. Thus, the first boxplot represents the range of timing errors occurring in all the first trials within each experimental condition; the second boxplot represents the range of timing errors within the second trial of each experimental condition; and so on for the third and fourth trials. The medians within the box-and-whisker plots for each trial are similar, suggesting that timing errors were generally similar across all four trials within each condition.

Figures 28-32 show no apparent systematic change in timing errors over trials, suggesting that no trends of improved or diminished performance were observed over the experimental session. Thus, trial order within experimental conditions and trial sequences over the experimental session may not be necessary as fixed effects in the model for timing error; however, the statistical analysis will determine whether trial order is needed.

4.6.4.3. Exploratory Plots of the Fixed Effects of the Experimental Conditions on Timing Errors

I will begin this section with exploratory box-and-whisker plots of the predictor variables of Feedback and Target Rate, and then show how the information in these plots was used in the statistical models for these data. In Figure 33, the timing errors for each experimental condition are displayed.

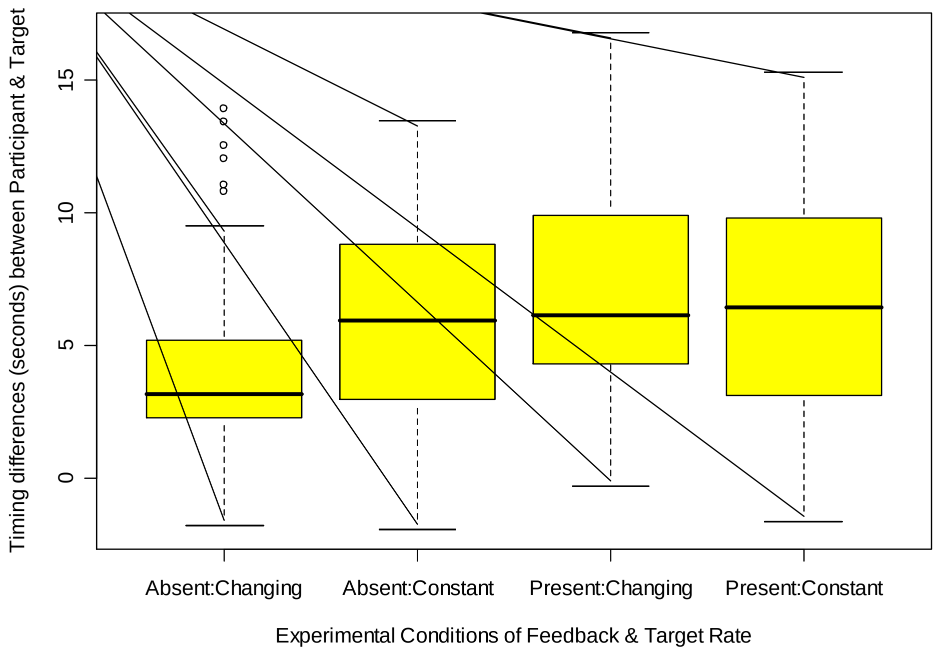

Figure 33. Timing errors for each experimental condition (*n* = 1430- 1536), measuring the differences between the participants and the target stimulus at the peaks within the four experimental conditions.

The box-and-whisker plots in Figure 32 show that the smallest timing differences between participants and targets occur in the Feedback:Absent-TargetRate:Changing condition. The boxplot for the Feedback:Absent-TargetRate:Changing condition has a lower median and a narrow range of values than the remaining experimental conditions (Feedback:Absent-TargetRate:Constant, Feedback:Present-TargetRate:Changing, and Feedback:Present-TargetRate:Constant). Since the differences in the location and spread of the values for timing accuracy across the four experimental conditions indicate that Feedback and Target Rate have different effects on the outcome they must be included in the initial model.

The boxplots in Figure 33 offer additional information about the timing errors under the four experimental conditions. These plots show that the majority of the timing errors between participant and target at corresponding landmarks were positive, meaning that the participants lagged behind the target at most of peaks in the data set. However, the lower whiskers in each experimental condition extend below zero, showing that each condition had several cases where a participant moved faster than the target for at least a portion of the trial. See Figure 36 for more details on timing differences at each peak.

Figure 34 reveals non-parallel lines across the levels of Feedback and Target Rate, indicating an interaction between the effects of Feedback and Target Rate on timing errors.

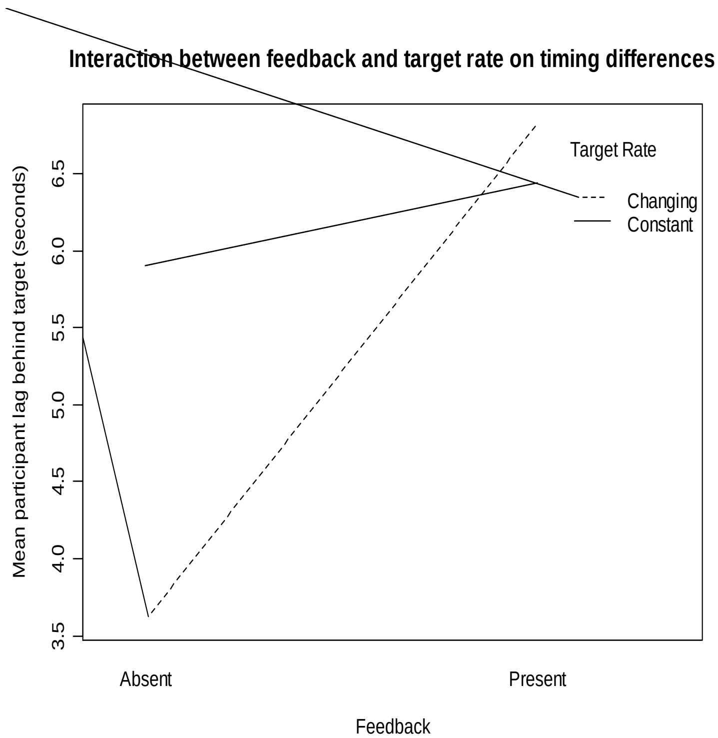

Figure 34. Interaction plot for Feedback and Target Rate on timing errors.

The crossed lines in Figure 34 indicate that the effect of the levels of Feedback (absent and present) depend on the level of Target Rate (changing and constant) and conversely, the effect of the Target Rate depends on the level of Feedback that is provided. Interactive effects make the interpretation of Feedback and Target Rate more complex. The presence of the interaction signals that the levels of Feedback and Target Rate are mutually dependent and cannot be interpreted as separate main effects (Underwood, 1997), as described in Section 4.5.5.2. Thus, Figure 34 specifies that the interactive term Feedback∗Target Rate should be included in the initial model. Due to a mathematical rule that requires terms to be entered into a model hierarchically, both Feedback and Target Rate will be included as main effects in the initial model, but they will not be interpreted separately.

4.6.4.4. Exploratory Plots of the Effect of Gender on Timing Errors

In Figure 35, the effect of the third predictor, Gender, is displayed in a lattice plot with the effect of Gender presented separately in each experimental condition.

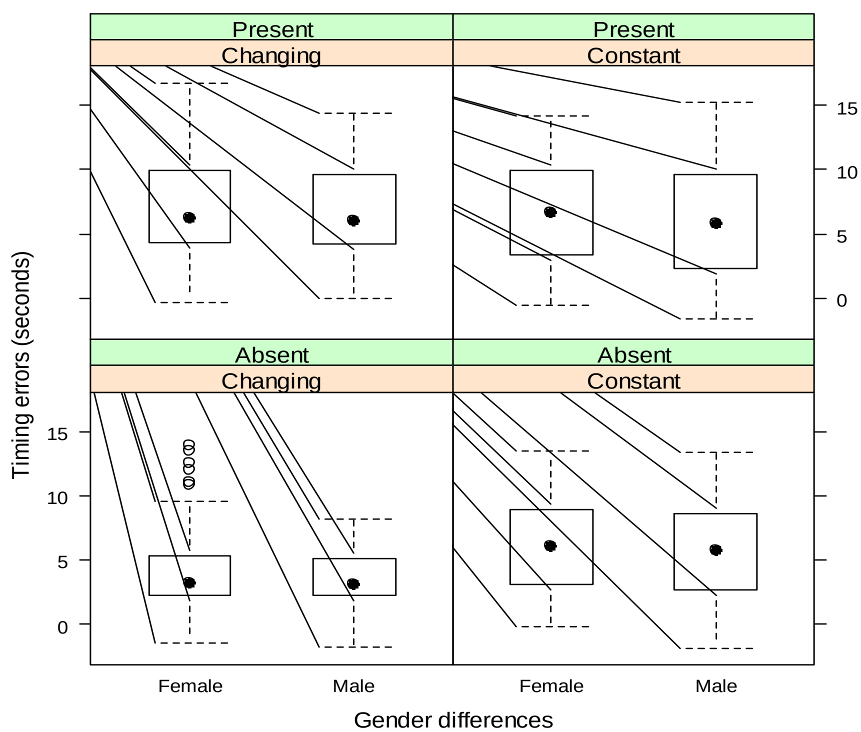

Figure 35. Timing errors by Experimental Condition and Gender (n: males = 470-528, n: females = 957-1010 responses in each of their respective boxplots). Within each box representing one of the four experimental conditions (Feedback:Absent-

TargetRate:Changing, Feedback:Absent-TargetRate:Constant, Feedback:Present-TargetRate:Changing, and Feedback:Present-TargetRate:Constant), the responses are divided into responses for females and males.

Figure 35 shows that the effect of Gender is smallest in the Feedback:Absent-TargetRate:Changing condition for both females and males, where the median response and spread of timing errors are very similar. Likewise, the median timing errors and spread of values for timing errors in females and males are similar to each other in the Feedback:Present-TargetRate:Changing condition. In the Feedback:Absent-TargetRate:Constant and Feedback:Present-TargetRate:Constant conditions, median timing errors are slightly lower for males than for females. Differences between participants and target are mostly positive, indicating that participants lagged behind the target at most observations, although males had more negative values than females, indicating that at some points in the sample, their movements led the target instead of lagging behind. The differences between genders shown in Figure 35 suggest that Gender should be included in the model; since Gender has different effects in the four experimental conditions, it is appropriate to evaluated Gender in the initial model within a three-way interaction, Feedback∗Target Rate∗Gender.

4.6.4.5. Exploratory Plots of the Effect of the Peak Order on Timing Errors

After establishing that the explanatory factors of Feedback, Target Rate, and Gender and the Feedback∗Target Rate∗Gender interaction should be included in the model, the next step involves evaluating differences in timing errors at each sample peak, and comparing them across experimental conditions. In Figure 36, timing errors occurring within samples at successive peaks are presented, with errors categorized by peak order and treatment (i.e., experimental condition).

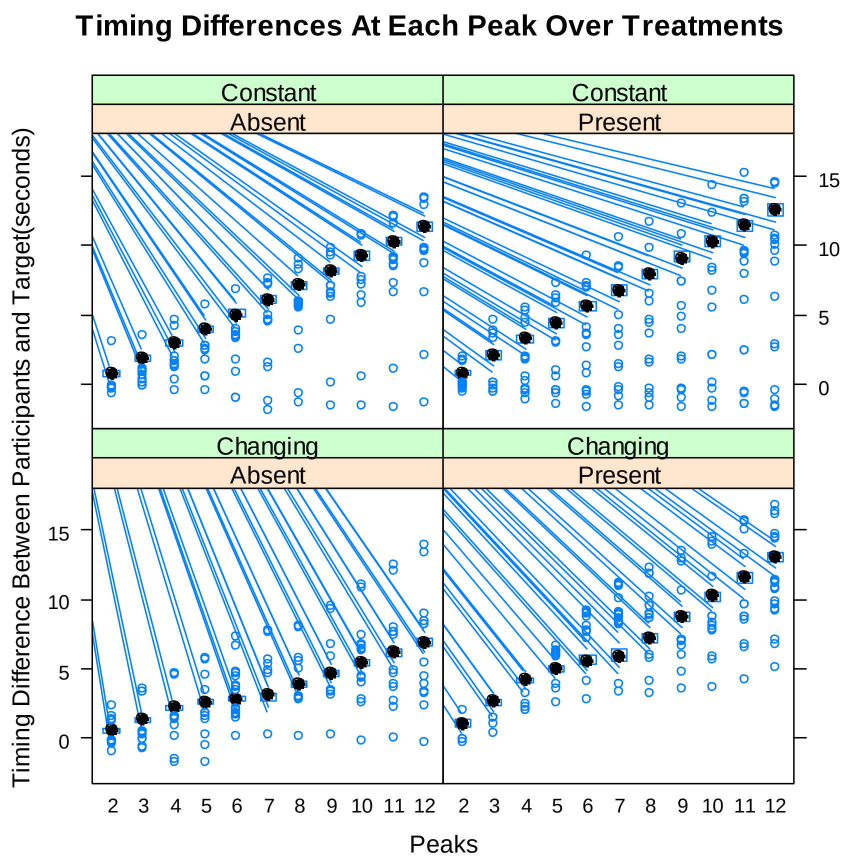

Figure 36. Timing errors (*n* = 5964) at each peak (1-12) in samples produced under each experimental condition. The black dots indicate the median values of the timing error between participant and target stimulus at each successive peak.

Although it is difficult to see the boxplots at each peak in this plot, the black dots are centered in box plots for each numbered peak; the box plots represent the central 50% of the values of timing errors at their respective peaks. The dashed horizontal lines slightly below and above the box plots indicate the boundaries of the first and third quartiles, respectively; and the blue dots represent outliers at each peak. The entire data set is represented here, including outliers.

Figure 36 reveals several interesting trends in the timing error data. First, most timing differences between participant and target movements (i.e., timing errors) showed that the participants lagged behind the target. In 5881 (98.6%) of the 5964 timing errors, participants fell behind the target. Second, the observed lags appeared to increase at each peak across the sample, with participants lagging behind the target increasingly at each successive peak. Third, a small percentage of timing measures (n = 83, or 1.4%) matched or exceeded the rate of the target, with some fast errors occurring in each condition, particularly in the Feedback:Absent-TargetRate:Constant condition and Feedback:Present-TargetRate:Constant condition.

Figure 36 demonstrates that the slopes that would be formed by connecting the black dots representing the medians of the timing errors at each successive peak are different for each treatment condition. The patterns shown by these exploratory plots indicate that most timing errors increased (i.e., produced increasing lag) systematically with each peak in the sample from Peak 2 to Peak 12 (since Peak 1 is the starting position and set to zero). Therefore, since the errors vary at each peak, Peak must be included in the initial model.

Figure 36 also shows that the data in the TargetRate: Changing conditions appears to be nonlinear. Although the timing errors in both TargetRate:Constant conditions (i.e., Feedback:Absent-TargetRate:Constant and Feedback:Present-TargetRate:Constant) increase linearly across the sample, the errors in the TargetRate:Changing conditions (i.e., Feedback:Absent-TargetRate:Changing and Feedback:Present-TargetRate:Changing), change slope at peaks 4-7. The change in slope in this area of the sample corresponds to the abrupt change in target rate over peaks 3-6 in the TargetRate: Changing conditions. The errors in the TargetRate: Changing conditions are linear before and after the timing perturbation (Peaks 1-2 and 7-12).

The presence of nonlinearity in the TargetRate: Changing condition is an important consideration for developing models to analyze these data. If non-constant variance is seen by the linear mixed effects model as well, an alternative model to the linear mixed effects model, such as the Generalized Additive Mixed Model, may be necessary to adequately fit these data.

4.6.4.6. Summary of Timing Data from Exploratory Plots

The exploratory plots in Figures 28-30 indicate that individual participants performed quite differently from one another, suggesting a linear mixed model with Feedback∗TargetRate∗Gender∗Peak as fixed effects, and with Participants as random effects. Variation in the timing errors across trials seemed to vary from participant to participant (i.e., some participants had a narrower ranges of variability than others, as shown in Figures 28-30). These observations provide additional support for the decision to model participants as random effects in the initial full (i.e., beyond-optimal) model for timing errors. Timing errors varied for each level of Feedback, Target Rate, Gender, and for Peak order, indicating that the beyond-optimal statistical model for these data must include each of these factors as fixed effects.

The patterns of the data shown in Figures 34-35 suggest an interaction among Feedback, Target Rate, and Gender (Feedback*TargetRate*Gender). This result indicates that the initial model for the timing errors in these samples must contain a three-way interaction among these factors. Thus, the initial linear mixed effects model will include fixed effects formed by a four-way interaction between Feedback, Target Rate, and Gender and Peak as fixed effects; random effects will include participants with trials (n = 16) nested in participants.

4.7. Applying Linear Mixed Model Methods to Timing Error Data

For the analysis of timing errors, the effect of concurrent visual feedback on timing errors, the stability of the feedback effect after a timing perturbation, gender, and peak location were explored. A model was developed after reviewing graphical displays of the data, including Figure 27, which showed the structure of the data, and Figures 34 - 36, which showed interactions among Feedback, Target Rate, Gender, and Peak.

The data structure diagram (Figure 27 in Section 4.6.3) shows that the random effects for the timing error data have a three-way structure: (a) peak are nested in their respective samples, (b) samples are nested within experimental conditions, and (c) samples from all experimental conditions are nested in participants. Section 4.3.2.2, *Expanding the Random Intercept Linear Mixed Model to a Three-Way Data Set,* discussed previously, described a three-way hierarchical model structure as Equation 4:

$$Y_{ijk} = \alpha + a_i + b_{j|i} + \varepsilon_{ijk} \qquad \text{(Equation 4)}$$

Applied to the timing error data, Equation 4 specifies an overall average for the timing error represented by α; average deviations from the overall timing errors for each participant represented by a_i; and an average deviation for each sample from the average for the corresponding participant represented by $b_{j|i}$ (i.e., the value of trial$_j$ given participant$_i$. The individual observations that deviate from the fitted timing error values at each peak produce the error terms, denoted by the term ε_{ijk}. Both a_i and $b_{j|i}$ are assumed to be normally distributed with mean 0 and variance σ_a^2 and σ_b^2, respectively. Differences between individual observations and the predicted timing error values at each peak in the corresponding sample produce the residual error after accounting for participant and trial effects, represented by ε_{ijk} ; they are specific to each observation (ijk_1 - ijk_n). ε_{ijk} is assumed to be normally distributed with mean 0 and variance σ^2. Note that this model specifies the hierarchical relationships among the observations but does not specify the experimental conditions and explanatory factors in this analysis.

The initial mixed effects model for timing was based on the graphical displays of the data and the underlying investigations: (a) the effect of vision on timing errors;

(b) the robustness of those effects with a timing perturbation (i.e., the TargetRate:Changing condition); (c) trends in those effects over time, shown by changing effects over the successive peaks in the sample; and (d) differences in performance between genders.

Thus, the initial timing error model contained these factors:

(1) Feedback, with values absent and present;

(2) Target Rate, with values changing and constant;

(3) Gender, with values female and male;

(4) Peak in the sample (2-12, after the starting position at peak 1).

The model also has to allow for the possibility that the explanatory variables may interact, meaning that the effect of any explanatory variable (Feedback, Target Rate, Gender, or Peak) on timing errors may depend on the level of another explanatory variable. For example, Figure 36 shows that the majority of timing errors increase with each successive peak, and that the slope of these changes differs for each experimental condition (Feedback:Absent-TargetRate:Changing; Feedback:Absent-TargetRate:Constant; Feedback:Present-TargetRate:Changing; Feedback:Present-TargetRate:Constant).

4.7.1. Step 1: Evaluating the Need for a Linear Mixed Model

After evaluating exploratory plots and building an initial model structure with the four-way interaction Feedback∗TargetRate∗Gender∗Peak to assess the data, West et al. (2007) and Zuur et al. (2009), advise that the next step in the analysis is determining whether the repeated measures data must be modeled with random effects, or if a fixed effects model would suffice. If the repeated observations were statistically the same, for example, a fixed effects model would eliminate the need to model participants as random effects. The procedure for determining whether a fixed effects model or a random effects model is needed to represent timing errors involves comparing these two models. Since the only difference between the fixed model and the linear mixed effects model is the random component, these models are nested and can be compared by calculating the likelihood ratio (Zuur et al., 2009) using ANOVA. The fixed model is denoted by Equation 17, *Timing-gls. Timing-lme* represents the linear mixed effects model with the same four-way interaction as fixed effects and random effects consisting of participants, denoted by $a_{i,}$ with trials nested in participants, denoted by $b_{j|I,}$ in Equation 18 *.*

Timing-gls: $Timing\,Error_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + \varepsilon_{ijk}$ (Equation 17)

vs.

Timing-lme: $Timing\,Error_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + a_i + b_{j|i} + \varepsilon_{ijk}$

(Equation 18)

The comparison of the fixed effects model and mixed effects model is displayed in Table 3.

Table 3

Comparison of Timing Data Using Generalized Least Squares and Linear Mixed Effects Models

	Model	df	AIC[a]	BIC[b]	logLik[c]	Test[d]	L.Ratio[e]	p-value[f]
Timing-gls	1	25	19214.63	19381.90	-9582.315			
Timing-lme	2	27	13746.18	13926.83	-6846.090	1 vs 2	5472.449	<.0001

The likelihood ratio statistic (L = 5472.449, p < .0001) shows that the fixed effects and the linear mixed effects models are significantly different. The two measures of model fit, the *AIC* (Akaike Information Constant) and *BIC* (Bayesian Information Constant) are lower for the linear mixed model, indicating that the linear mixed model provides a significantly better fit to these data than the fixed model structure. Thus, the mixed model structure was used for these data.

4.7.2. Step 2: Initial Modeling Procedures: Developing the Beyond-Optimal Model for Timing Errors

After determining that the mixed model structure was appropriate for the timing error data, I developed an initial model based on the research questions, the data structure, and the exploratory plots as discussed in Sections 4.6.3 and 4.6.4. Ten data points were removed as outliers. In this study, outliers are defined points that fall more than 1.5 times the interquartile range above the third quartile or 1.5 times the interquartile range below the first quartile (Renze, n.d.).

I began the analysis by applying a random intercept mixed effects model to the analysis of the timing error data. Recalling that Figure 36 indicated that timing errors in the TargetRate:Changing conditions were nonlinear, I was aware that this analysis might have required a more complex model, but chose to start with the simplest structure to see if issues of nonlinearity could be easily resolved and if the model would exhibit constant variance. If the model did not have these features, the next step would be to consider a random slope model, a random coefficient model with both random intercepts and random slopes, or a generalized additive mixed model.

The first step in model development for mixed models involves setting the structure for the fixed effects. Instead of using mathematical notation, I am writing the fixed terms of the model in words, after Galwey, (2006):

$$\text{Feedback} * \text{Target Rate} * \text{Gender} * \text{Peak} + \text{Trial Number}_{\text{conditions}} + \text{Trial Number}_{\text{session}} \quad \text{(Formula 11)}$$

The * symbols in Formula 11 indicate an interaction. In this instance, the term Feedback * Target Rate * Gender * Peak means that all main terms, two-way interactions, three-way interactions, and the four-way interaction among the main terms are included in this model. The four-way interaction in this model accounts for changes in each explanatory variable at each level of the other explanatory variable. Since terms enter the model hierarchically, a mathematical rule requires that all three-way, two-way and main terms be included when a four-way interaction is present.

The random portion of the linear mixed effects model includes an intercept for participants (a_i) to allow us to estimate differences between participants (intersubject variability), and a random intercept for trials ($b_{j|i}$) for estimating differences between trials by the same participant (intrasubject variability), nested in the corresponding participants.

Combining the fixed part (Formula 11) and random part of the model, and applying them to the factors in this study, we have the initial linear mixed effects model in Equation 19:

$$Timing\ Error_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Peak_{ijk} * Gender_i + Trial_{ab} + a_i + b_{j|i} + \varepsilon_{ijk} \quad \text{(Equation 19)}$$

The index system *ijk* now stands for observation *k* in trial *j* from subject *i*. The term a_i is an intercept specific for each subject, and the $b_{j|i}$ is an intercept for each trial j. This notation appears complicated, but it allows each subject to have higher or lower values than others (a_i), and each trial to have higher or lower values than others by the same participant (via $b_{j|i}$). In addition, it measures the response at each level of feedback, target, and peak for each participant, allowing each participant to respond to every condition. Since each participant can have only one gender, a single subscript is used to represent the gender category for each participant. Similar models can be seen in Pinheiro and Bates (2004, p. 43) and West et al. (2006, p. 227).

The linear mixed effects model in Equation 19 indicates that Timing $Error_{ijk}$ refers to each observation, and $Feedback_{ijk}$, Target $Rate_{ijk}$, $Gender_{ijk}$, and Peak $_{ijk}$ refer to the effect of Feedback, Target Rate, Gender, and Peak on timing errors, allowing each participant to have separate intercepts (i.e., overall means, allowing intersubject variability), and each trial within participants to have separate intercepts (i.e., trial means, allowing intrasubject variability). Trial order$_{ab}$ is evaluated separately to determine if trends in performance occurred over the trials in each experimental condition, *a* (1-4) or over the experimental session, *b* (1-16). Thus, the analysis of this model will assess timing differences between participant and target at each peak, and evaluate the interaction of Feedback, Target Rate, and Peak on the timing differences, allowing for separate effects for males and females. The random components are reflected in the term a_i, which allows separate intercept values for each participant; $b_{i|j}$, which denotes trials nested in participants; and ε_{ijk} which refers to deviations from the predicted timing errors at each peak in each trial nested in the respective participant.

4.7.2 1. Initial Modeling Procedures: Checking For Order Effects in Timing Error Data.

Beginning the model development process, I assessed the effects of trial order to determine whether the data depended on the order of the four trials within each experimental condition or the sixteen randomized trials in the experimental session (termed trial sequence in this analysis). Thus the first model that I analyzed included the four-way interaction among Feedback, Target Rate, Gender, and Peak, with trial order within each experimental condition and trial sequence over the experimental session added separately as fixed effects (Equation 19). Participants comprised the random effects in this model.

Trial order within each experimental condition, *a* (1-4) was not found to be significant statistically (F=. 81, p = .49) in the linear mixed model defined by Equation 19. Trial sequence over the set of trials over the experimental session, *b* (1-16) was nonsignificant statistically as well (F= .033, p = .855). Thus, trial order within experimental conditions and trial sequence within the experimental session were not included as fixed effects (i.e., predictors) in the first tested model, allowing trial sequence to be nested in participants as a repeated measure.

Since trial orders in each experimental condition and across the experimental session, were not found to be significant, the terms of the *beyond-optimal* linear mixed effects model (Zuur et al., 2007) included the four-way interaction between Feedback, Target Rate, Gender, and Peak order. Participants comprised the random effects in the model, with trial sequence nested in Participant to enable an assessment of intrasubject variability. The model was analyzed with REML (restricted maximum likelihood). Thus, the protocol for model fitting was begun with the model denoted by Equation 20:

$$Timing\ error_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Peak_{ijk} * Gender_{ijk} + a_i + b_{j|i} + \varepsilon_{ijk}$$

(Equation 20)

In this model, feedback, target rate, peak, and gender are fixed effects. Random effects include participants and trials nested in participants. Together, these variables comprise a random intercept linear mixed effects model, where both participants and trials within participants have separate intercepts, but the same slope.

4.7.3. Step 3: Initial Modeling Procedures: Checking For Constant Variance

Once the model has been selected, the next step in model development involves plotting and assessing the model residuals for constant variance. In the timing error data, the residuals from the fully loaded (i.e., beyond-optimal) model denoted in Equation 20 displayed structure that varied with increasing values of the predictors, as shown by Figure 37.

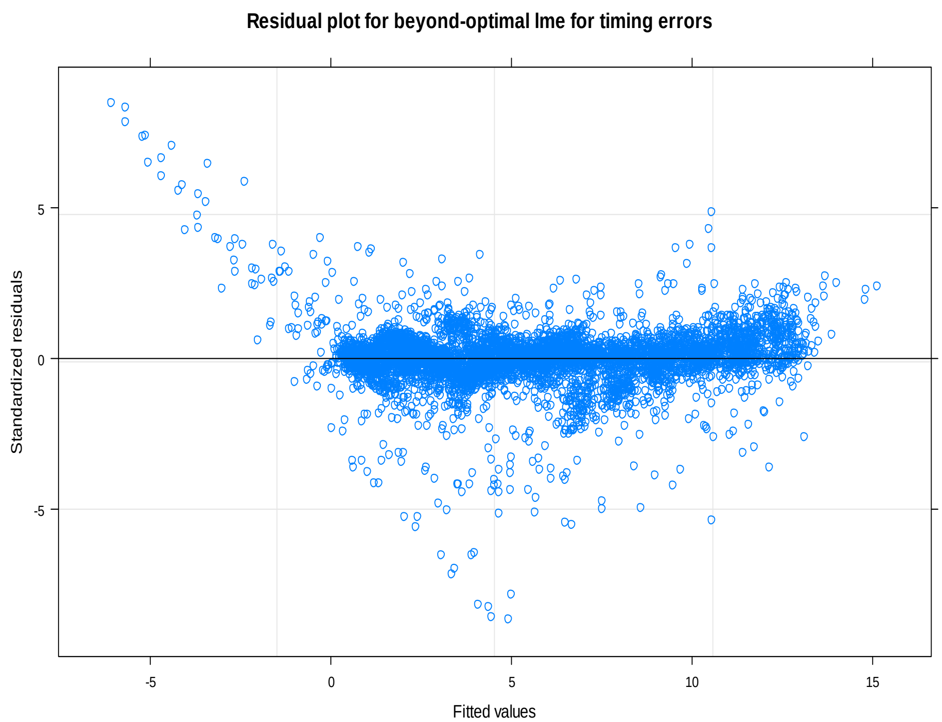

Figure 37. Standardized residuals plotted against fitted values of timing errors for initial full model (Equation 20) for the timing error data.

Since standardized residuals have variance = 1, a standardized residual that is larger than two is considered large. In Figure 36, the standardized residuals range from approximately -8 to 8, approximately, indicating that the model is not adequately specifying a substantial part of the data set. Further, the residuals for the initial model denoted by Equation 20 showed a clear systematic pattern. Even though the linear mixed model was an improvement over the fixed effect model, patterned residuals indicate that a simple linear mixed model did not adequately account for the structure in the data. Additional exploratory plots are needed to examine the data to reveal features of the data that may be contributing to the patterns in the residuals.

Recalling the nonlinearity in Figure 36, I plotted the residuals against peak and applied a smoother to the residuals, as shown in Figure 38. The variability appeared to be associated with peaks 3-7, which corresponded to the peaks where the Target Rate varied in the changing Target Rate condition.

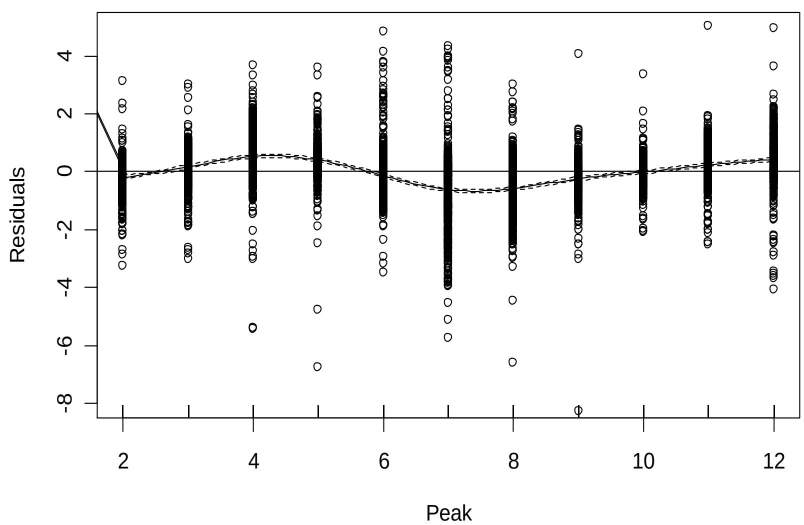

Figure 38. A plot of the model residuals versus peak from the *beyond-optimal* linear mixed effects model (Equation 20) for timing error.

A smoother was added to visualize patterns in the residuals. The results of Figure 38 prompted a review of the residuals under the two Target Rate conditions (changing and constant), as shown in Figure 39 and 40.

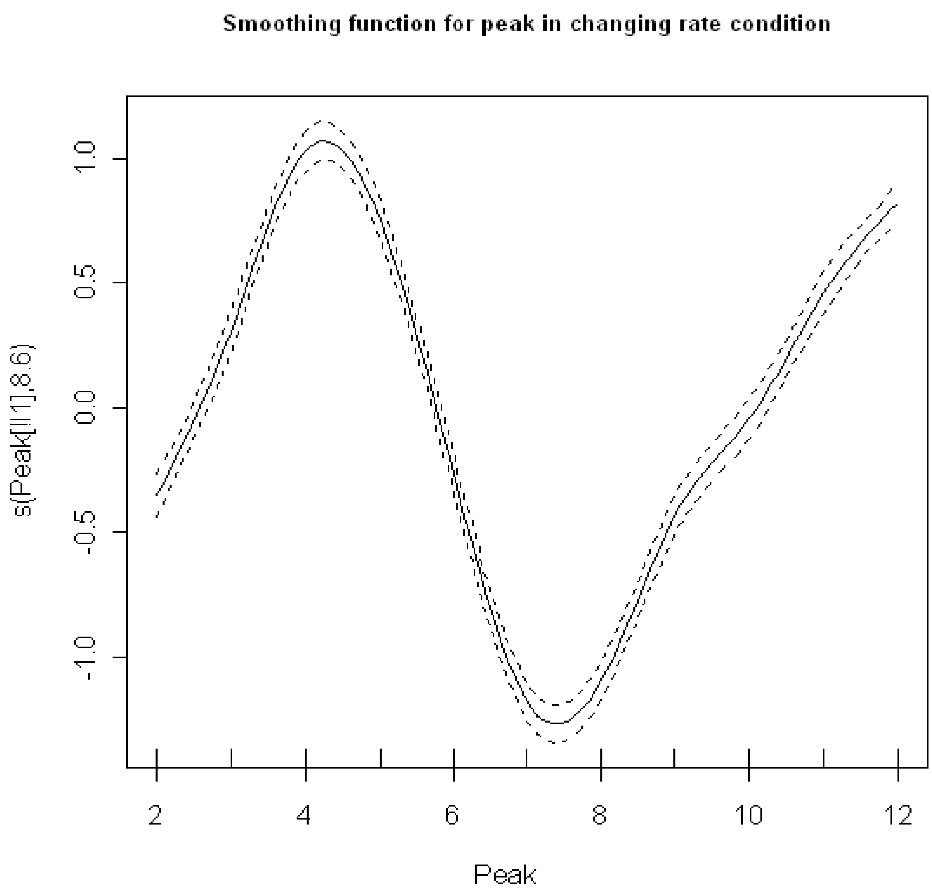

Figure 39. Residual patterns of timing errors for changing Target Rates.

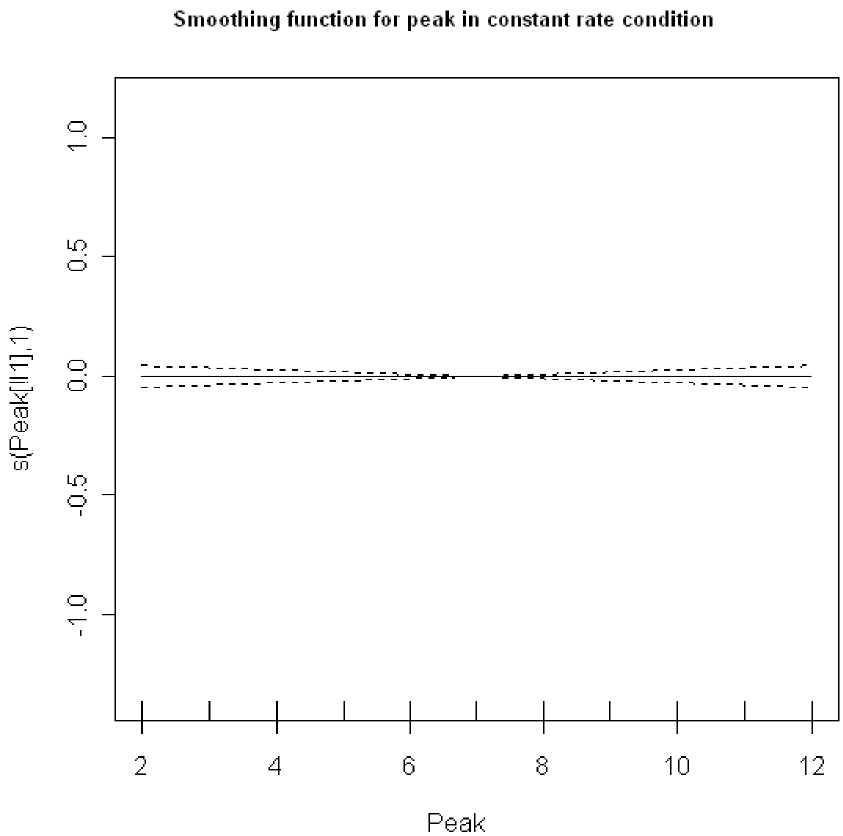

Figure 40. Residual patterns of timing errors for constant Target Rates.

In Figure 39, note that residuals at peak 3, peak 4, and peak 5 are mostly positive, and are mostly negative at peak 6, peak 7, and peak 8, as can be seen from the curve produced by the smoother. These patterns result when a mixed model imposes a linear relationship on data that are nonlinear. The smoother function imposes a value of 8.6, shown in the y-axis in the plot of residuals vs. peak in the TargetRate:Changing condition. Since smoother values that deviate from one indicate nonlinearity, the value of 8.6 on the *y*-axis in Figure 39 indicates that the data in the TargetRate:Changing condition are considerably nonlinear. In Figure 40 for the TargetRate:Constant data, the smoother imposes a value of one (1), shown in the y-axis in the plot of residuals vs. peak in Figure 40, indicating that those data have a nearly perfectly linear relationship with the predictors.

Figure 41 shows the patterns of residuals for each experimental condition. Note that the smoother is nearly linear in the TargetRate:Constant conditions (upper plots), but is considerably nonlinear in the TargetRate:Changing conditions (lower plots). The presence or absence of feedback appears to make little difference for the value of the smoothers in these plots: the greater influence is the difference in Target Rate (changing or constant).

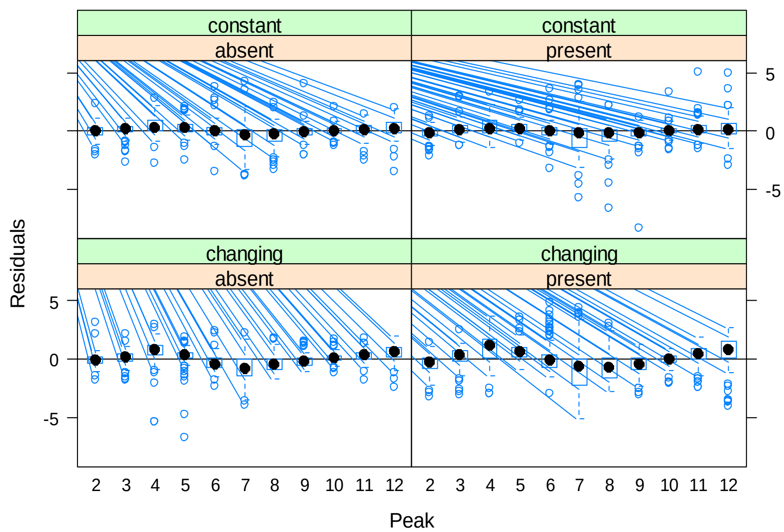

Figure 41. Pattern of residuals in linear mixed model framework for timing error, categorized by treatment type, denoted by the four quadrants in the lattice plot. The black dots indicate the medians of the boxplots for timing errors at each peak in the samples (2-12, on the lower axis). The dashed lines represent the *whiskers*, the extent of the data from the 25th-75th percentile. The blue open circles represent outliers.

The residual pattern in Figure 38 and the nonlinearity seen in the residual plots (Figure 39 and Figure 41) indicate that the current linear mixed model is not adequate for these data. The shape of the patterns in the residuals indicates that the present beyond-optimal model described in Equation 20 will not produce reliable statistical tests. The waves in the residual patterns also suggest that extending the beyond-optimal model with quadratic terms is not likely to produce a linear model whose residuals are scattered randomly across zero. Transformations are another strategy that can sometimes reduce non-constant variance in the residuals, but transformations are not advised when the residuals reflect an important aspect of the data (Zuur et al., 2007). Further, since the TargetRate:Changing conditions provide valuable insight into the patterns of participants' responses, transformations are a less attractive option to solving the problems of nonlinearity in these data. Thus, adding a random slope to the model may be a more appropriate method of handling the nonlinearity in these timing error data.

At this point in the analysis of the timing error data, several facts are known: a mixed model is necessary to model the random participant effects, but a random intercept model was not sufficient for dealing with the nonlinearity and non-constant variance in the data. Further, exploratory plots in Figures 38 - 41 showed that predictor variables yield different responses at different peaks (i.e., a peak effect), so the model for these data must account for changing responses as the participants advance from peak to peak across the sample.

As an alternative, in modeling data that do not show consistent patterns is to allow each set of factors to have its own trajectory (intercept and slope) which show the unique trends for each combination of feedback, target rate, and gender over the 11 peaks in the sample. Suppose, in the current data, that the relationship between feedback and target rate is different in each combination of factors. A model that explicitly accounts for differences in variation for each combination at different points across a sample will be most appropriate. Such a model, which measures a sample at repeated intervals while allowing the slopes of each combination to vary is termed a random slope model, and provides information about differences across experimental conditions (between-condition data) and within each experimental condition. Since these data reflected this type of situation, the random slope model may be an appropriate choice for these data..

4.7.4. Step 4: Building a Random Slope Mixed Model for the Timing Error Data

4.7.4.1. Selecting Fixed Effects for a Random Slope Model for Timing Error Data

Recalling that the model in the present study is an attempt to assess the fixed effects of Feedback, Target Rate, Gender, and Peak order on the timing errors at successive peaks, the model begins with all possible fixed effects and their interactions, termed the *beyond-optimal* model (Zuur et al., 2009). Peak order was entered as a linear variable since using Peak order as an 11-level variable was too costly in terms of degrees of freedom and plots of the data indicated that timing errors (lags) increased linearly with each peak. Since trial orders were not significant, the fixed effects are expressed as

Feedback * Target Rate * Gender * Peak order (Formula 12)

Due to the interaction, peak order was centered in the model. Centering allows easier interpretation of the values of the terms in the model, and reduces the likelihood of multicollinearity. Since the two levels of each of the other predictor variables (Feedback, Target Rate, and Gender) are coded as 0 and 1, respectively, centering is not necessary. The fixed portion of the model including Feedback, Target Rate, Gender, and the centered values for Peak is shown in Formula 13. Note that Formula 13 depicts fixed effects only, and does not include the random component for the random slopes model.

$$TimingError_{ijk} = \alpha + Feedback_{ij} * TargetRate_{ij} * Gender_{ij} * Peak(centered)_{ij}$$

(Formula 13)

The indices *ijk* denote the *k*th observation in the *j*th trial ($j_{1\text{-}16}$) which is nested in the corresponding *i*th participant, measuring the difference between a participant and the target at a peak in the sample.

4.7.4.2. Selecting a Structure for Random Factors

Once the fixed effects are determined, the random effects are selected to handle correlations in the data. A bit of experimentation may be required to find the best random effects structure. A lower AIC was obtained with the random effect of trial sequence nested in participants, allowing for participant-specific random slopes across the peaks in the sample. Peaks were centered to aid in interpretation.

Thus, the *beyond-optimal* model*,* containing all possible combinations of predictors with the selected random effects is expressed by Equation 21.

$$TimingError_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * \ Gender_{ijk} + Peak(centered) + a_{i1} + a_{i2} + b_{j|i} + \varepsilon_{ijk}$$

(Equation 21)

The indices *ijk* denote the *k*th observation in the *j*th trial ($j_{1\text{-}16}$) which is nested in the corresponding *i*th participant, measuring the difference between a participant and the target at a peak in the sample. The fixed effect terms consist of the three-way interaction Feedback∗TargetRate∗Gender plus the centered Peak term. The random effect terms are designated by a_{i1} and a_{i2} which allow for separate intercepts and slopes for participants (a_i), respectively, with trials nested in participants ($b_{j|i}$). The residuals, ε_{ijk}, are obtained from the differences between the *k*th observation and the corresponding fitted values for the observations in each sample, nested in the corresponding participant.

4.7.4.3. Evaluate Initial Model for Constant Variance

Once a potential model is selected and applied to the data, the next step is to determine whether the strategies for dealing with nonlinearity and non-constant variance are effective. The *beyond-optimal model with random slopes* shown by Equation 21 yielded the residual pattern shown in Figure 42.

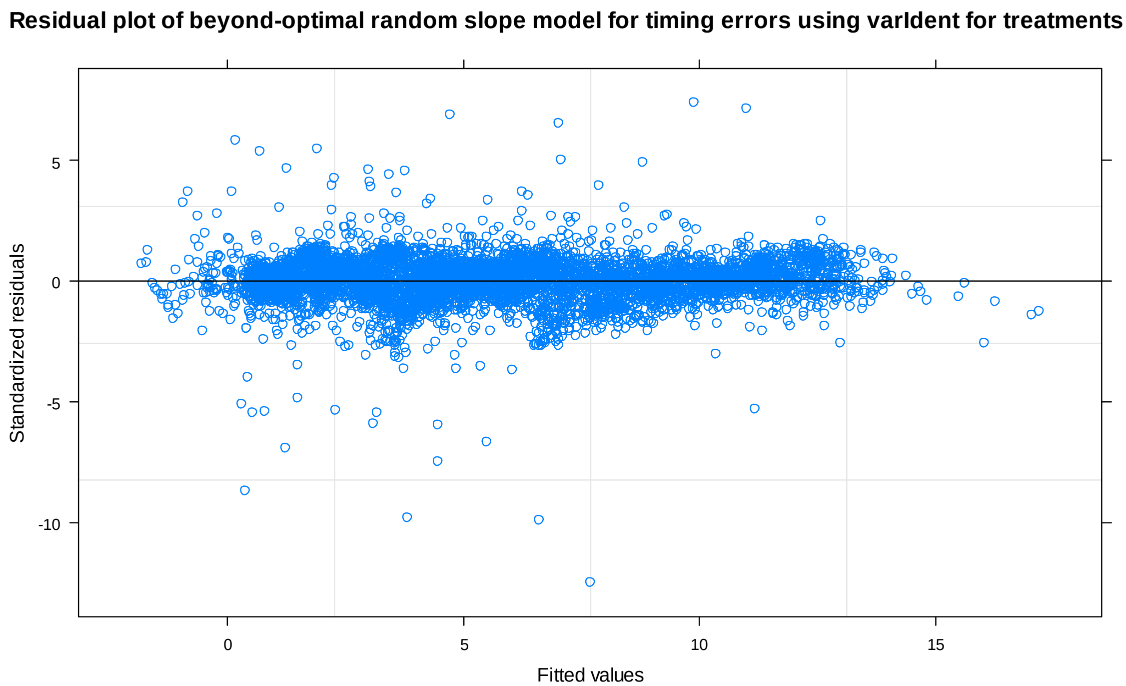

Figure 42. Residual patterns of beyond-optimal model (Equation 21), using a model that included a random slope for peak along with trials nested in participants as random effects.

Although the quantile-quantile plot for these data was judged mildly non-normal by its shape and size of the data set, the impact of non-normality on the inference was considered to be limited (Faraway, personal communication, July 12, 2008). Further, Tamhane and Dunlop (2000) claim that the normality assumption is not critical for very large sample sizes. The residuals are distributed across zero with a reasonably random structure (Zuur, personal communication, March 18, 2009). Thus, we can assume that no additional manipulations of the covariance structure are needed, using the unstructured covariance structure, which is the default covariance structure in *R*.

4.7.4.4. The Effect of Trial Sequence over the Experimental Session in the Random Slope Model

Following the initial model building, trial sequence in the experimental session (trial 1-16) was plotted against model residuals to assess for any effects of trial sequence in the experimental session, shown in Figure 43. With the exception of some outliers, the values for the residuals of the *beyond-optimal* random slope model were concentrated around zero for all sixteen trials, indicating that trial sequence did not appear to have a significant effect on the timing differences, and verifying the earlier statistical tests of trial order reported in Section 4.7.1.

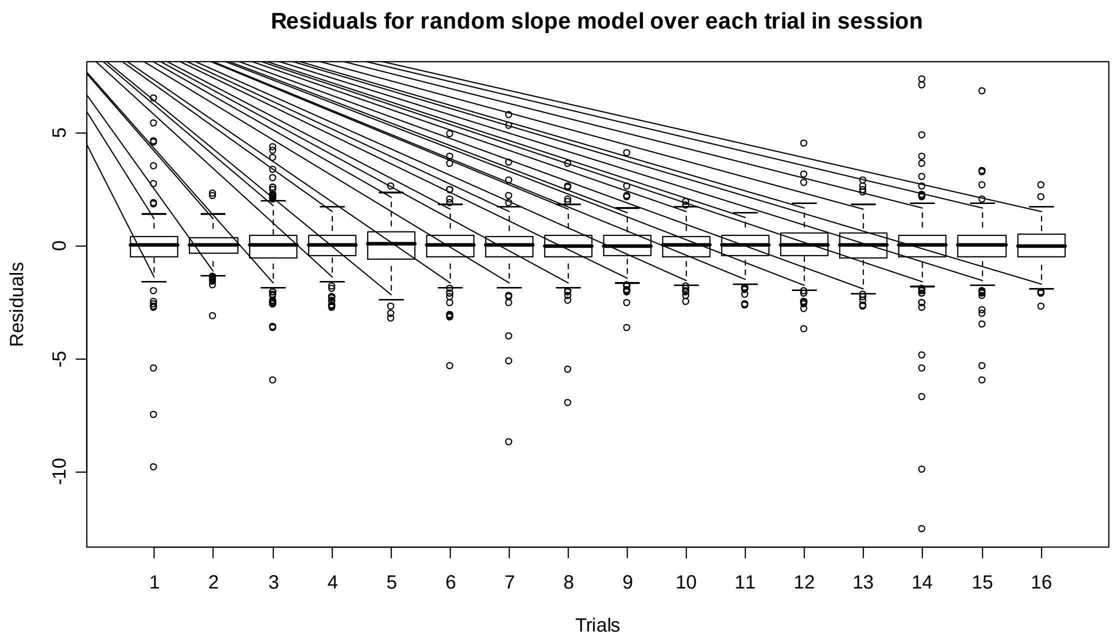

Figure 43. Residuals of timing error data for each sample in the experimental session.

The plot of the residuals of the random slope model against the trial numbers in the experimental session shows that the means for each trial do not vary much from each other. This indicates that individual participants were relatively consistent across trials in their timing errors.

4.7.5. Step 5: Refining the Model for Timing Error Data

Following the model selection procedures outlined in West et al. (2007) and Zuur et al. (2009) which use the Akaike Information Criterion (*AIC*) as the standard, I dropped statistically nonsignificant terms from the *beyond-optimal model* sequentially using stepwise backward selection, based on p-values. Beginning with the beyond-optimal model that included Feedback, Target Rate, Gender, the centered Peak, and all possible interactions among these variables, the largest non-significant term was removed at each step. Because of the large sample size, non-significant terms were defined as those having a *p*-value of 0.01 or greater.

When model simplification was complete the optimal model was given by terms in Equation 22:

$$\begin{aligned}
&TimingError_{ijk} = Feedback_{ijk}*TargetRate_{ijk}*Peak(centered)_{ijk} + Feedback_{ijk}*Gender_{ijk}*Peak(centered)_{ijk} \\
&+ Feedback_{ijk}*TargetRate_{ijk} + Feedback_{ijk}*Gender_{ijk} + Feedback_{ijk}*Peak(centered)_{ijk} \\
&+ TargetRate_{ijk}*Peak(centered)_{ijk} + Gender_{ijk}*Peak(centered)_{ijk} \\
&+ Feedback_{ijk} + Target_{ijk} + Gender_{ijk} + Peak(centered)_{ijk} + \varepsilon_{ijk} \\
&\text{with varIdent covariance structure for } Feedback_{ijk}*TargetRate_{ijk}
\end{aligned}$$

(Equation 22)

Since the final reduced model can be difficult to understand from the equation, the *R* syntax for this model is shown in Formula 14, and includes the effects of the explanatory variables *Feedback, Target Rate, Gender, and Peak,* and significant interactions between these terms on the response, *Timing Error*. *Trials* nested in *participants* enter the model as trial-specific intercepts, and *centered peaks* enter the model as random slopes. Feedback, TargetRate, and Gender are factors (2 levels each), while Peak is treated as a continuous variable because of the cost in terms of degrees of freedom.

TimingErrors<-lme(Time.Difference ~

Feedback*TargetRate*Peak + Feedback*Gender*Peak

+ Feedback*TargetRate + Feedback*Gender+ Feedback*Peak

    + TargetRate*Peak + Gender*Peak

    + Feedback + TargetRate + Gender + Peak

Fixed effects

+ random= ~Peak|Subject/Trial, method="REML",

+   weights=varIdent(form=~1|Treatment)

Random effects

Formula 15

4.7.6. Step 6: Final Tests of the Timing Error Model

4.7.6.1. Assessment of Residuals

After the final model is attained, it is appropriate to perform some final tests of residuals to validate the model to ensure that it provides a suitable fit to the data.

Figure 44. Plot of standardized residuals vs. fitted values of the optimal random slope model for timing errors, as depicted by Equation 22. Trial-specific random intercepts are included in the model by default.

The plot of the final timing error random slope model shows that the residuals are distributed across zero. While the variances do not appear to increase or decrease systematically with the means and the structure appears reasonably random, some structure remains that was not accounted for in the present model. Further exploration is warranted, but for the purposes of the present study, the distribution of residuals from the final random slope model appears to satisfy the minimal requirements for randomness. To learn more about the patterns of residuals from the final optimal timing model, I performed additional analyses by subdividing the data set by the four experimental conditions, shown in Figure 45.

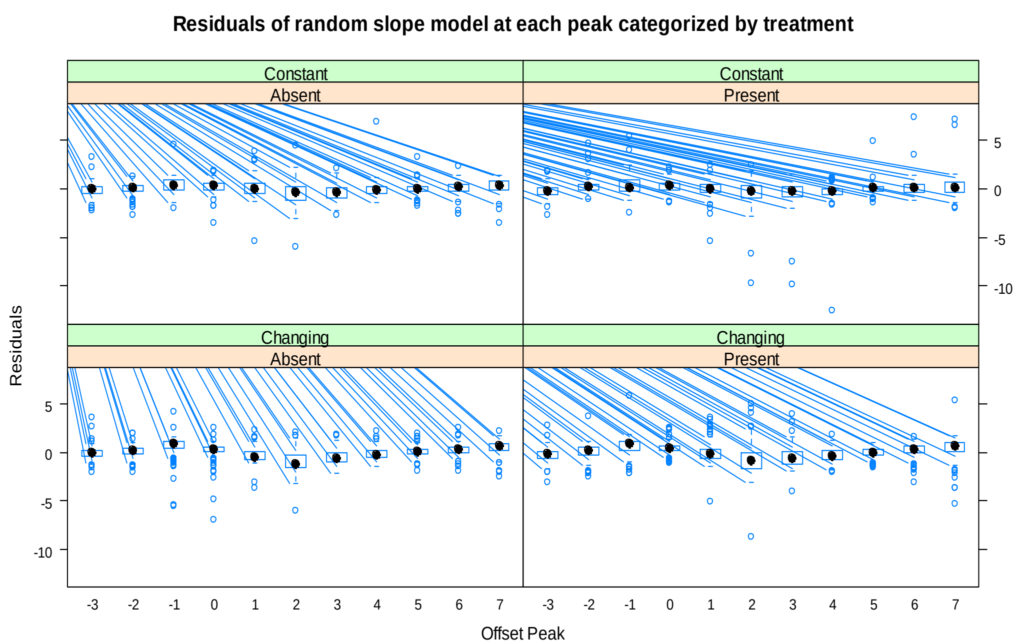

Figure 45. Residuals of the optimal random slope model for timing errors as depicted by Equation 22, categorized by treatment.

The black dots in Figure 45 represent the median values for the residuals from the optimal random slope model for the timing error (i.e., Equation 22). The medians for the optimal model are clustered much more closely around zero than the residuals in the random intercept mixed model without random slopes or the varIdent structure represented in Figure 41. Although they are very hard to see, the medians are located in the center of boxplots for the timing data, which contain the central 50% of the data. The placement of the medians close to zero indicates that majority of residuals do not increase or decrease at successive peaks in the samples, although some heterogeneity continues to be present. Dashed lines in Figure 45 represent the boundary marking the lower 25% and upper 75% of the data. The open blue dots represent outliers that extend beyond those boundaries.

A smoother placed through the data in each condition suggests that normality has been considerably improved using random slopes for Peak and allowing each level of treatment to have its own variance (varIdent structure), as shown in Figure 46.

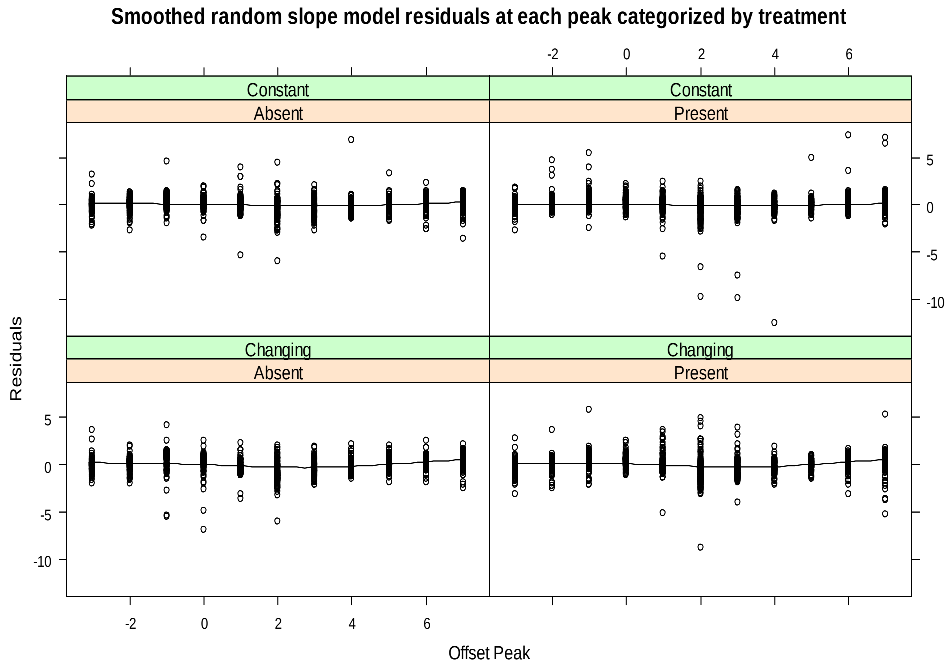

Figure 46. Smoother placed through residuals to evaluate normality of residuals in random slope model for timing errors, as depicted by Equation 22.

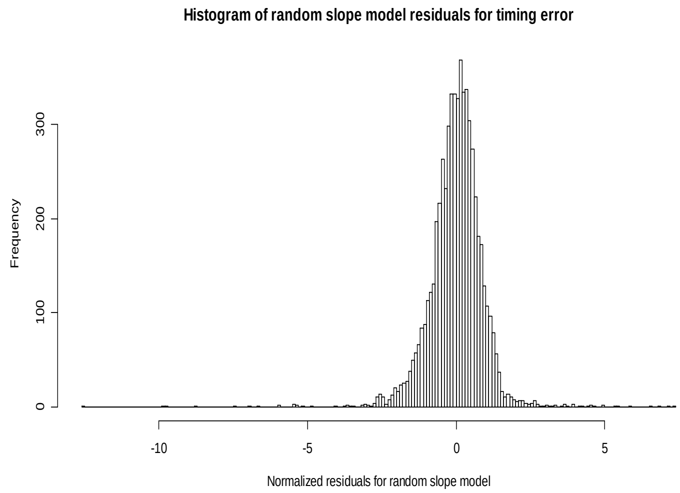

Figure 47. A histogram of residuals for the optimal random slope model for timing errors as depicted by Equation 22.

Figure 47 shows that residuals for the optimal model produce a relatively normal histogram with some outliers and a slight left skew. The overall normality of this histogram provides additional evidence that the random slope provided a reasonable approach to the nonlinearity in the timing error data.

4.7.6.2. Interactions

In Section 4.5.5.2, I reported that interactions pose a difficulty for interpreting the effect of an explanatory variable within the interaction. Statisticians such as Underwood (1997) and Fox (2002) contend that it is not possible to interpret main effects independently of the interaction in which they occur. In rare cases, limited interpretations are possible (Oehlert, personal communication, June 5, 2009). By using likelihood ratio comparisons in deletion tests, I explored whether the main effects of Feedback and Target Rate had statistically significant effects on the timing differences by nesting models with and without these terms. In *Chapter 5, Results,* I show that models that lack one of the predictors, Feedback and Target Rate, have significantly poorer fits to the data. Thus, each substantially contributes to a better understanding of the data.

4.7.7. Final Comments on the Random Slope Model for Timing Error Analysis

In summary, the final model (Equation 22), reproduced here, appears to fit the timing error data adequately.

$$\begin{aligned}
&\mathit{TimingError}_{ijk} = \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk}*\mathit{Peak(centered)}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Gender}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Gender}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{TargetRate}_{ijk}*\mathit{Peak(centered)}_{ijk} + \mathit{Gender}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{Feedback}_{ijk} + \mathit{Target}_{ijk} + \mathit{Peak(centered)}_{ijk} + \varepsilon_{ijk} \\
&\text{with varIdent covariance structure for } \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk}
\end{aligned}$$

(Equation 22)

The residual plots in Figures 44-46 and the histogram of the residuals in Figure 47 suggest that the most effective model of those tested allowed random slopes for peaks, participant specific intercepts for trials, and varIdent structure for treatments which allow different variances for each level of treatment. With this model, heteroscedasticity was reduced and data appeared reasonably normal. Thus, a random slope modelling procedure appeared to be an adequate choice for analyzing timing errors, and effectively accounted for the subject-specific effects, correlated observations within trials and participants, heteroscedasticity, and nonlinearity that characterized the timing error data.

4.8 Exploration of Spatial Differences between Target and Sample Drawings

4.8.1. Characteristics of the Spatial Error Data

The term *spatial error data* in the present study refers to tracing errors defined as the difference between the shapes of the curves drawn by the target and by the participant for each peak-to-peak segment across the sample (n = 11). Spatial errors are measured as the root mean square error (RMSE) between each corresponding curve between target and participant. This approach yields a vector of 11 responses, each representing a RMSE value for the corresponding peak-to-peak segment, as described in Section 3.4 in Chapter 3, *Segments Selected for Comparison and Analysis.* If each participant contributed sixteen samples, complete with 11 peak-to-peak curves drawn in each sample, 176 (11 x 16) root mean square error (RMSE) scores would be added to the data set. Of 6160 possible RMSE values in the data set from 35 participants (35 x 176 scores), 5976 RMSE values were used in the analysis. Scores from 176 curves were unavailable due to missing samples caused by researcher error, equipment malfunction, or participant error. In addition, eight outliers were removed from the final data set.

The spatial errors (i.e., RMSE scores) were expected to form sets of correlated observations as follows:

(1) Spatial errors at each of the eleven peak-to-peak curves within a sample were correlated with each other.

(2) Spatial errors of each participant were correlated within one another in each experimental condition.

(3) Spatial errors of each participant are correlated with one another across the entire set of sixteen trials contributed by each participant.

The data structure depicting these correlations for the spatial error data is a three-way nested data structure, with responses nested in trials, and trials nested in participants.

4.8.2. The Starting Point: Exploratory Data Analysis With Graphical Displays

As discussed in the wiggle section (4.4.2) and timing section (4.6.2), graphical assessments are the first steps in data analysis, since they enable us to assess the structure of the data. Exploratory plots enable us to visualize nonlinearity, lack of independence in the data, or serious violations of non-normality (Oehlert, 2000) that would require special consideration in the model development phase of the analysis.

4.8.3. Data Structure Diagram

Figure 48 displays a hierarchical three-level structure for the data, relating each measurement to other measurements in the data set. For clarity, only one participant is depicted in Figure 48, and only two trials are depicted with their respective peaks, under two conditions: one under Feedback:Absent-TargetRate:Constant and the second under the Feedback:Present-TargetRate:Constant conditions. To preserve readability and space, trials in the Feedback:Absent-TargetRate:Changing and the Feedback:Present-TargetRate:Changing conditions are not expanded here.

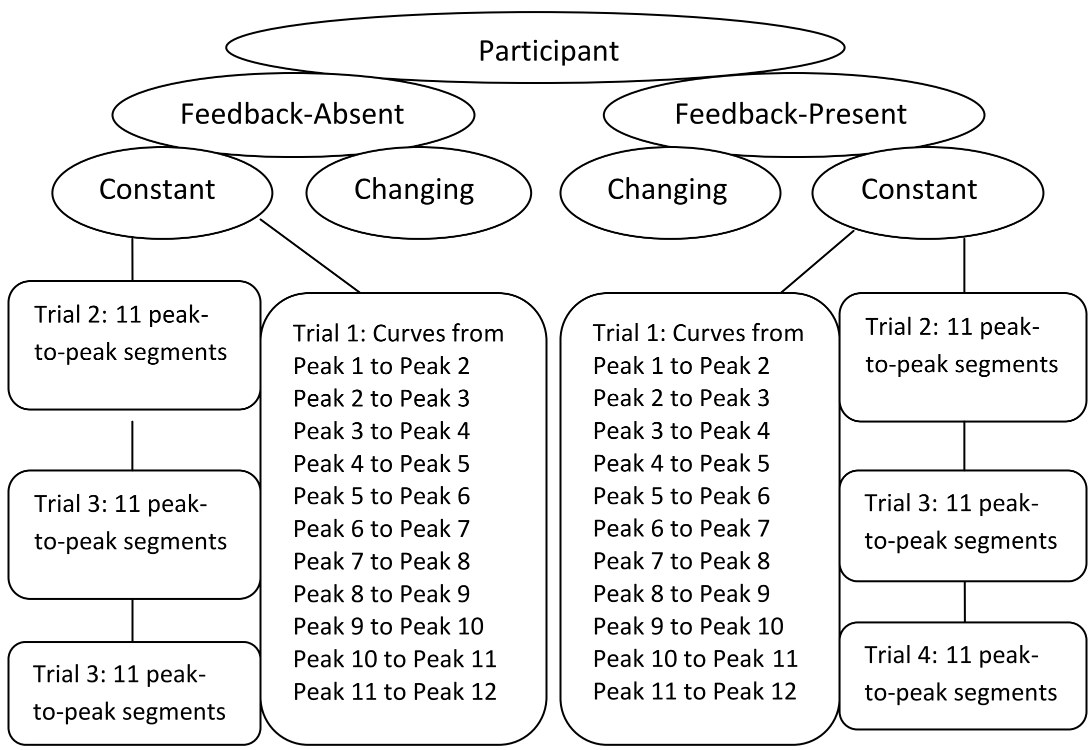

Figure 48. Arrangement of 11 peak-to-peak segments comprising each sample nested in each trial.

Figure 48 shows two trials (one Feedback:Absent-TargetRate:Constant and one Feedback:Present-TargetRate:Constant) expanded to show the eleven peak-to-peak segments nested in each trial. The remaining TargetRate:Constant trials are denoted by boxes with the words "Trial *n*: Times from peak 1 to peak….". Trials performed in the TargetRate:Changing conditions (four with Feedback-absent and four with Feedback-present) are not depicted here. The Target Rate conditions are placed under each Feedback condition to indicate that the four experimental conditions result from crossing two levels of Feedback (absent and present) with two levels of Target Rate (changing and constant).

The data diagram in Figure 48 displays three sources of variation for each participant:

(1) The first source of variation arises from RMSEs in the eleven measurements observed from each trial (one for each peak-to-peak segment in the sample).

(2) The second source of variation is the differences in RMSEs by each experimental condition.

(3) The third source reflects differences in observations within participants at corresponding landmarks.

This data diagram shows the dependence of measures by demonstrating that observations are correlated within trials, within experimental conditions (i.e., combinations of Feedback and Target Rate), and within each participant. In each spatial error data set, a fourth source of variation arises from the differences among individual participants, but to preserve clarity and space in Figure 48 the data structure for only one participant is shown. Only two trials are depicted with their respective peak-to-peak segments, one under Feedback:Absent-TargetRate:Constant and the second under the Feedback:Present-TargetRate:Constant conditions.

4.8.4. Exploratory Plots: Spatial Data

In this section, I will begin by displaying the spatial errors (RMSE scores) by participants measured under the four experimental conditions in this study. Next, I will display plots to view trends across samples and trials. Then I will display plots of the participant RMSE scores under each experimental condition. Each of these graphical displays will contribute to decisions made in the model building process.

4.8.4.1. Exploratory Plots of the Participant Spatial Errors

Figure 49 displays the spatial errors (i.e., RMSE scores) by participants (n = 35).

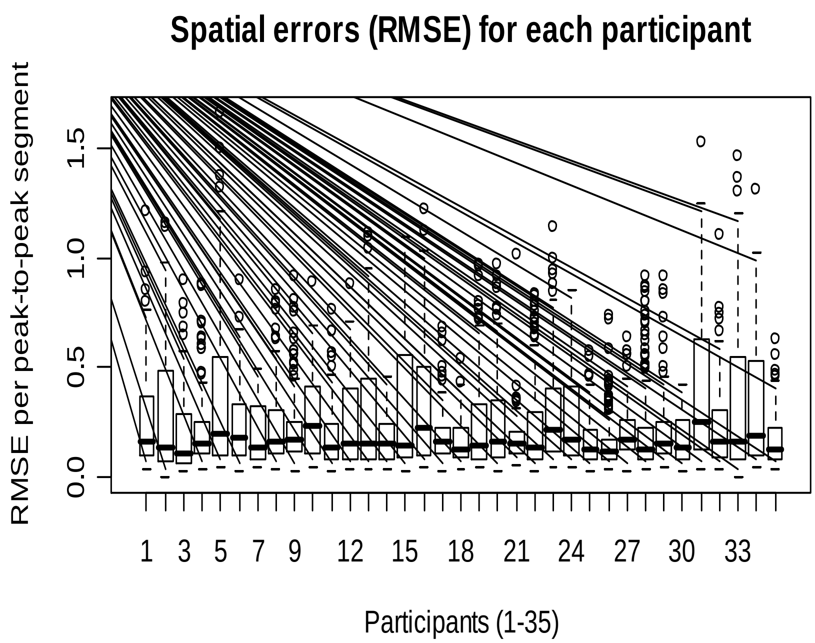

Figure 49. RMSE values for each peak-to-peak segment by participant. Each boxplot represents all the RMSE value measures for each participant (n = 158-176).

In Figure 49, the dark horizontal bars in the boxplots represent the median RMSE for the corresponding participant. Each box represents the central 50% of the data. The upper and lower whiskers, represented here by hatched lines, extend to the upper 25% and lower 75% of the data, respectively. The open black dots are outliers. Figure 49 shows that each participant contributed considerably different RMSE scores from another. In addition, the variability of RMSE scores within many of the participants varies widely. In later plots, we will see whether the variability occurred primarily in the Feedback:Absent condition where the participants were unable to monitor their drawing performance, or whether other conditions contributed to the variability within and between participants.

The patterns of spatial errors by each participant in the sixteen trials in the study are displayed in Figure 50. Each participant plot contains 16 box-and-whisker plots, one for each trial performed by a participant. The black points shown in the center of the boxplots represent the median RMSE for a participant from the 11 RMSE scores collected in each trial. Unfortunately, many of the individual boxplots are not visible. The open blue dots represent outliers.

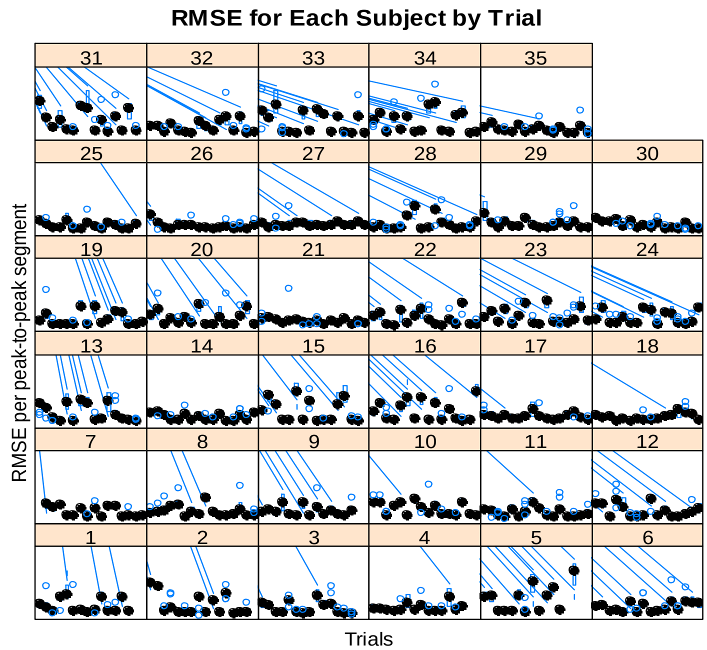

Figure 50. RMSE measures for each peak-to-peak segment (*n* = 9 - 11) in each trial (14-16) performed by each participant (*n* = 35).

Figure 50 reveals that some participants were relatively consistent over all trials, while others showed considerable variability. Given the small size of the box-and-whisker plots in the lattice plot, it is difficult to make a conclusion about trend in these data. To explore trend further in a clearer display, I have presented the RMSE measures for a subset of six randomly selected participants in Figure 51.

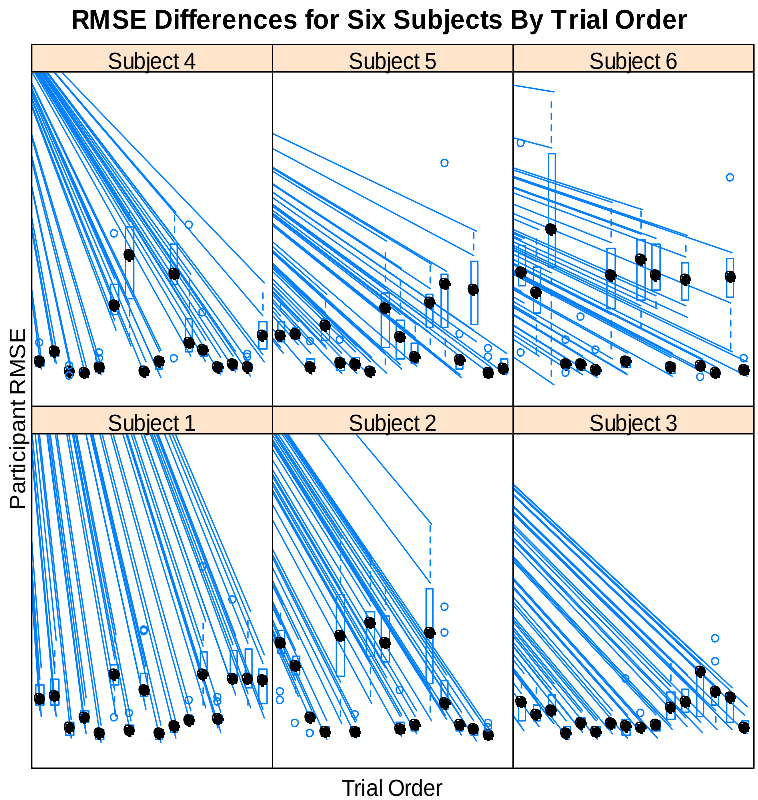

Figure 51. RMSE values from six randomly chosen participants, represented by the separate lattice plots, each containing box-and-whisker plot (*n* = 16), representing trials 1-16 contributed by each participant.

Figure 51 shows that each participant displayed here produced an idiosyncratic pattern of responses. The plots for five participants (Participant 1, 2, 4, 5, and 6) revealed considerable variation across trials; the boxplots for Participant 3 show a narrower range of variation. The medians for the individual trials for each participant are different in Figures 49, 50, and 51, and justify the decision to test participants as a random effect in the model. Since trials performed by each participant vary among each other, as seen in Figure 50 and 51, the model for spatial accuracy should include trials nested in the corresponding participant as random effects in the model.

4.8.4.2. Exploratory Plots of the Fixed Effects of Trial Order on RMSE

In addition to determining whether trends appear across trials for each participant viewed separately, it is important to assess for the presence of trend across trials in sequence (1-16) within the experimental session, as shown in Figure 52.

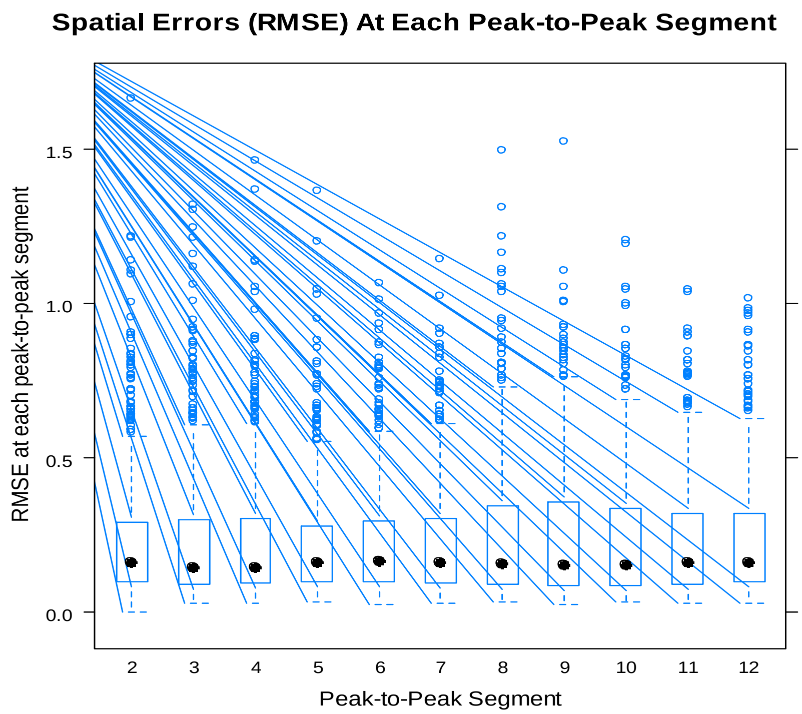

Figure 52. RMSE values contributed by all participants (n = 35) at each peak-to-peak segment (n = 11, from 2-12).

The view shown in Figure 52 displays the range of RMSE responses at each successive peak-to-peak segment in the samples produced by the participants in the present study. The median RMSE value for the segments (peak 1-peak 2, peak 2-peak 3, …, peak 11-peak 12) performed by all participants is indicated by the black dot in the box-and-whisker plot. The box-and-whisker plots show no systematic improvement or decline over the segments in each sample, suggesting that generally similar errors were made for all segments during a trial.

Figure 53 provides another view of trend by displaying the range and median of RMSE values occurring in each trial (trial 1, trial 2, …, trial 16) in the experimental session across all participants. Figure 53 indicates that few changes were found in the median RSME values for each trial after the initial orientation trials (trial 1 and trial 2). Thus, performance did not appear to improve or decline greatly over the randomized trials in the experimental session.

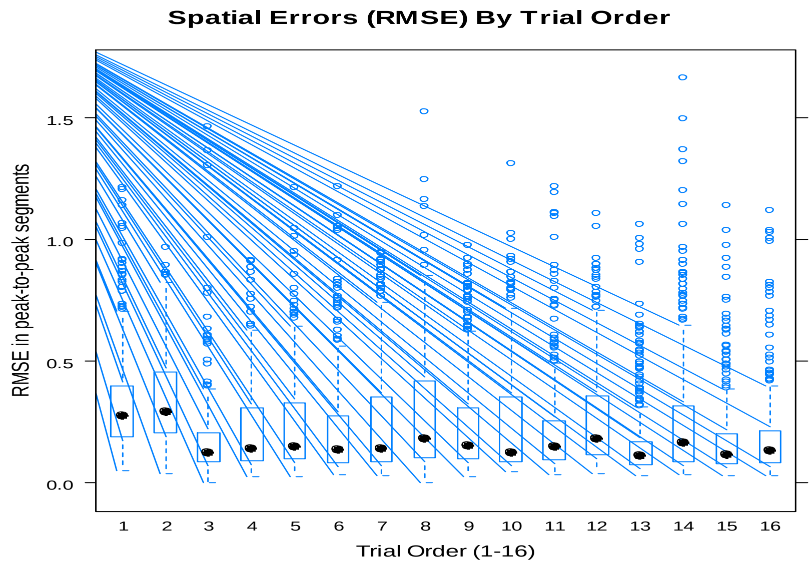

Figure 53. Box-and-whisker plots of RMSE measures from all participants (*n* = 35) at each trial (1-16) during the experimental session.

Each box-and-whisker plot in Figure 53 represents the range of responses occurring by trial order (trial 1, trial2, trial 3, and trial 4) under each experimental condition. Since the trials were presented to participants in a randomized manner, however, it is possible that the randomization masks systematic patterns within experimental conditions. In Figure 54, RMSE values are viewed according to the order of the trials within each experimental condition.

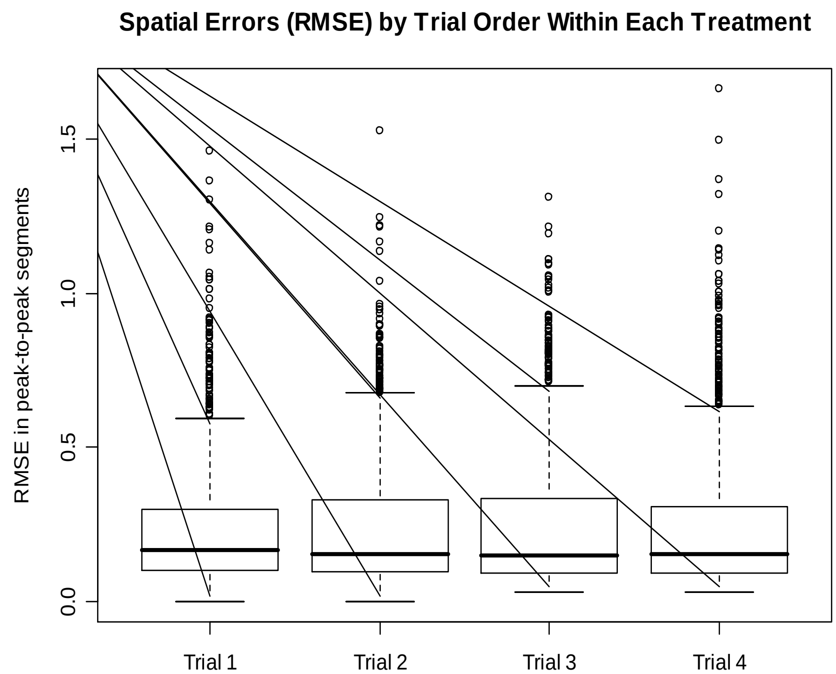

Figure 54. Root mean square error (RMSE) values at each trial (1- 4) within all experimental conditions.

The median RMSE values at each trial performed in sequence during the experimental conditions (i.e., trial 1, trial2, trial 3, and trial 4) produced similar medians, suggesting that RMSE scores were generally similar across the four trials performed in each condition. Since Figures 53 and 54 show no systematic change in RMSE measures over trials, we may observe no trends of improved or diminished performance are apparent in these plots. Therefore, it is unlikely that performance improved or declined over the experimental session. Nevertheless, this hypothesis will be assessed numerically in the model-building phase.

4.8.4.3. Exploratory Plots of the Fixed Effects of the Experimental Conditions (Feedback and Target Rate) and Gender on Spatial Errors

*Effects of Feedback * Target Rate on spatial errors.* I will begin this section with exploratory box-and-whisker plots of the experimental conditions, and then show how the information in these plots was used to develop the statistical models for these data.

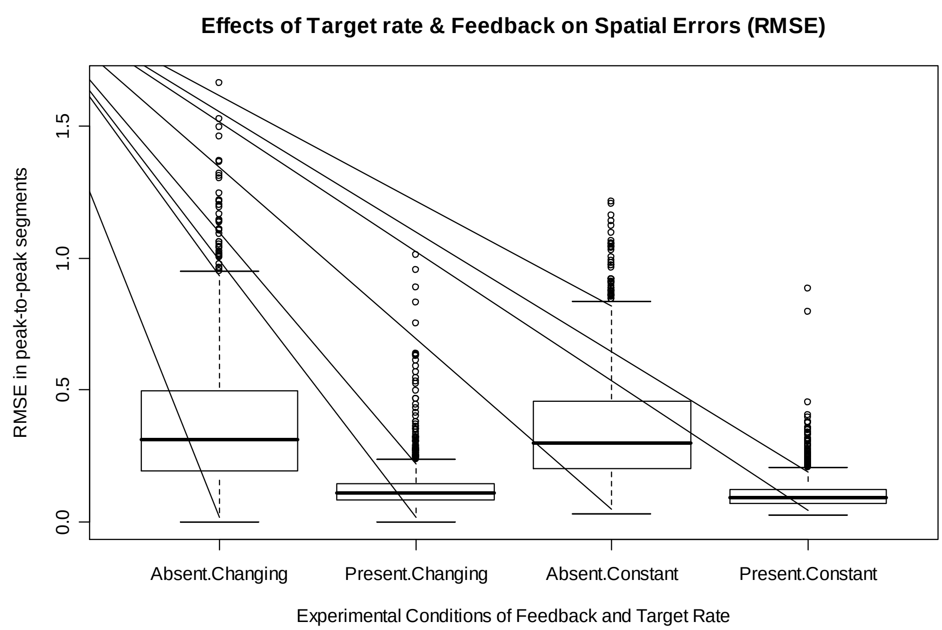

Figure 55. RMSE measures at each experimental condition (i.e., Feedback:Absent-TargetRate:Changing; Feedback:Present-TargetRate:Changing; Feedback:Absent-TargetRate:Constant; and Feedback:Present-TargetRate:Constant

conditions). The horizontal bar in the center of each box-and-whisker plot represents the median RMSE in each experimental condition. The edges of the boxes enclose the central 50% of the data (i.e., the interquartile range); the edges of the boxes indicate 25^{th} (lower edge) and 75^{th} (upper edge) percentiles, respectively. Vertical dashed lines extend 1.5 times the interquartile range in both directions, which is roughly two standard deviations above or below the mean (Crawley, 2007). The open black circles at the ends of the whiskers represent outliers.

In Figure 55, the RMSE measures presented in the box-and-whisker plots show that the smallest RMSE values and narrowest range of errors occurred in the Feedback:Present conditions (i.e., Feedback:Present-TargetRate:Changing and Feedback:Present-TargetRate:Constant), while larger RMSE values were found in the Feedback:Absent conditions. The differences in the location and spread of the values for RMSE measures across the four experimental conditions indicate that the factors comprising the experimental conditions (Feedback, Target Rate, and Feedback∗Target Rate) must be included in the initial model.

Effects of Gender on spatial errors. Differences in RMSE values by Gender under each of the four experimental conditions are depicted in Figure 56.

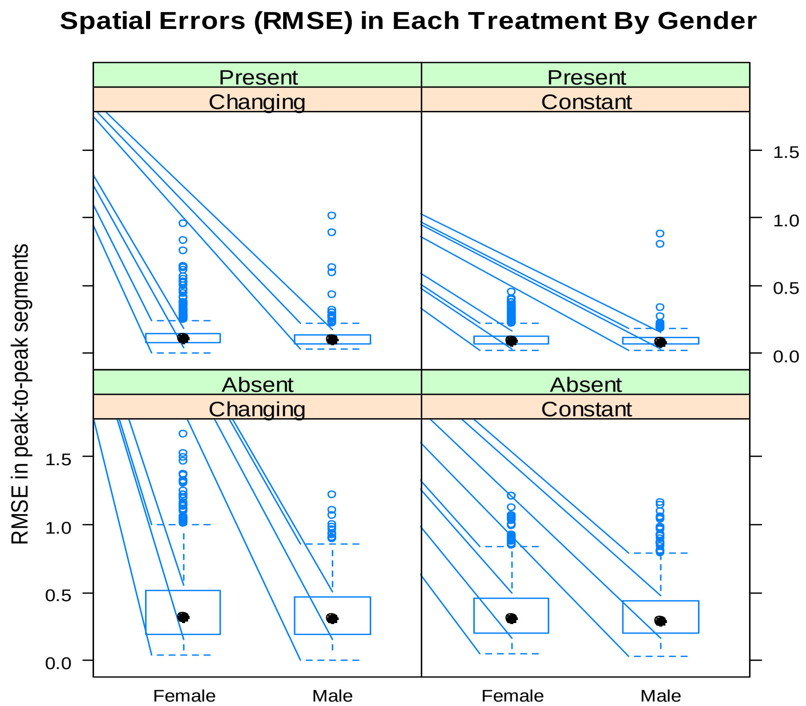

Figure 56. RMSE of spatial errors for females (n = 23) and males (n = 12) divided by experimental condition (Feedback and Target Rate, represented by the four quadrants).

The data in each box in Figure 56, representing an experimental condition, are subdivided further into two categories by gender, displayed as separate box-and-whisker plots in each quadrant. Dark points in each boxplot in Figure 56 indicate median values. The edges of the boxes indicate 25th (lower edge) and 75th (upper edge) percentiles, respectively. Vertical dashed lines extend 1.5 times the interquartile range in both directions, which is roughly two standard deviations above or below the mean (Crawley, 2007). Points above the vertical dashed lines are considered extreme values. Differences in RMSE values between male and female participants appear to exist between the Feedback:Absent and Feedback:Present conditions, but few gender differences are evident in the TargetRate:Changing and TargetRate:Constant conditions. Although differences between males and females on RMSE values appear to be small, Gender will be included in the initial model to assess its effect numerically on RMSE values, keeping in mind that the small sample size is underpowered for a reliable analysis. Thus, the three-way interaction Feedback∗TargetRate∗Gender will be used as a fixed factor in this analysis.

4.8.4.4. Exploratory Plots of the Effect of Peak-to-Peak Segments on Spatial Errors

After establishing that the explanatory factors of Feedback, Target Rate, and Gender and the Feedback∗Target Rate∗Gender interaction should be included in the model, the next step involves evaluating differences in spatial errors at each sample peak, and comparing them across experimental conditions.

Figure 57 shows details about the size and spread of the RMSE values within each experimental condition at each Peak-to-Peak segment. Each box-and-whisker plot represents the RMSE values for Peak-to-Peak segments extending from peak 1-2, peak

2-3, peak 3-4, …,peak 11-12 over the four trials performed in the experimental condition indicated by the respective quadrant of the lattice plot.

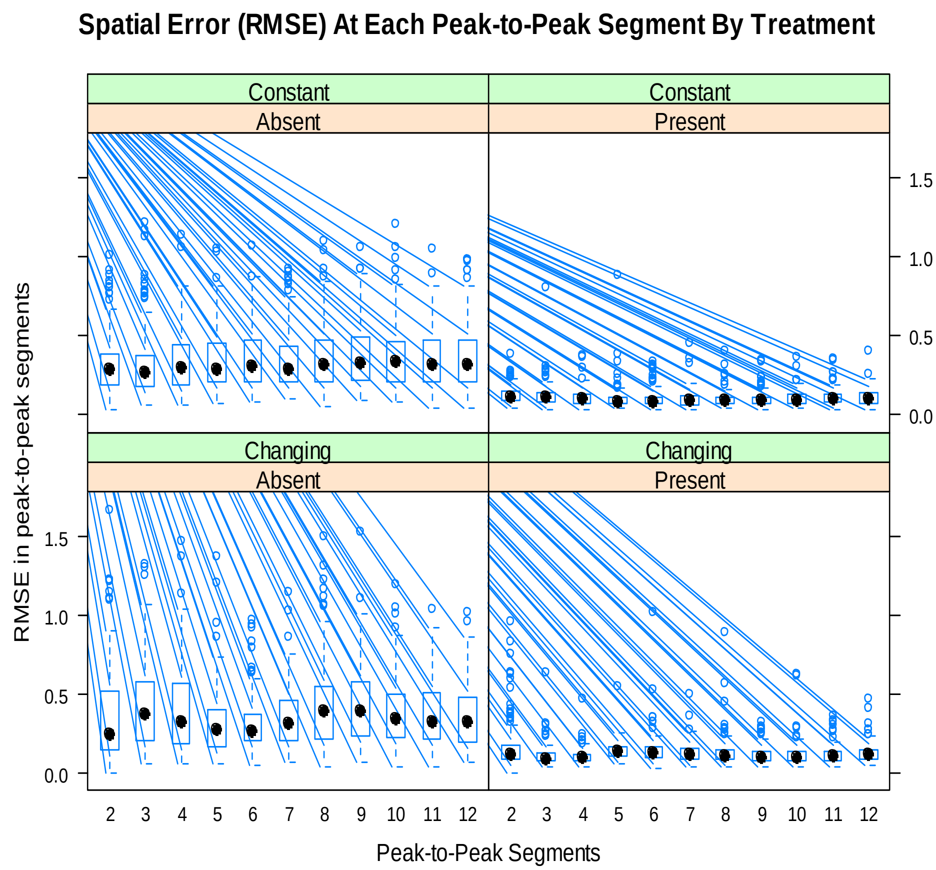

Figure 57. RMSE values at each peak-to-peak segment in the four experimental conditions represented by the four quadrants of the lattice plot.

Figure 57 reveals several interesting observations in the spatial error (RMSE) data. First, RMSEs between participant and target movements were higher in the Feedback:Absent conditions, as one might expect since participants cannot see their response to the tracking stimulus. Variability in the Feedback:Absent conditions was also broader than in the Feedback:Present conditions. A slight wave is evident in the TargetRate:Changing conditions near Peak-to-Peak segments 3-7, corresponding to the timing perturbation that occurred in the TargetRate:Changing condition. Overall, however, the median RMSE values for each Peak-to-Peak segment are quite consistent, suggesting that spatial errors neither increased nor decreased across the sample. Since the RMSE values varied in each experimental condition, a term for *Peak-to-Peak Segment* will be included as a predictor variable in the initial model.

4.8.4.5. Summary of Findings of RMSE Data from Exploratory Plots

The exploratory plots in Figures 49, 50, and 51 indicate that individual participants performed quite differently from one another, suggesting that a mixed model with participants as random effects may be appropriate. Variation across trials produced by participants seemed to be vary from participant to participant (i.e., some participants had a narrow range of variability, while others had much broader ranges). This finding supports the decision to nest trials in participants as a random effect in the initial model. The first choice of model for these data will be a linear mixed effects model, but if further analysis determines that a linear mixed effects model is inadequate for these data, a generalized additive mixed model may be a useful alternative.

The variability of RMSE values among the four experimental conditions indicates that both Feedback and Target Rate should be included as fixed effects in the initial model, but the effects of Gender are less clear. The effects of the order of the Peak-to-Peak Segments in the sample (peak 1-2, peak 2-3, …, peak 11-12) may also have an effect on the response. To assess their effects, Peak-to-Peak Segment and Gender will be included in an initial ("loaded" or *beyond-optimal*) linear mixed effects model as fixed effects. Thus, the initial linear mixed effects model will include fixed effects formed by a four-way interaction between Feedback, Target Rate, Gender, and Peak-to-Peak segment (I.e., Feedback∗Target Rate∗Gender∗PeakSegment; random effects will include Participants, with trials (n = 16) nested in participants.

4.9 Application of Linear Mixed Model Methods to Spatial Error Data

For the analysis of spatial copying errors (RMSE), the effect of visual Feedback on RMSE, and the stability of the Feedback effect when a timing perturbation occurred was investigated. In addition, the effect of Gender, and peak location (Peak-to-Peak Segment) was assessed for their effects on the observed RMSE values.

A model with these factors as fixed effects was developed after reviewing graphical displays of the data, including Figure 48, which showed the structure of the data, and Figures 54-56, which showed interactions among Feedback, Target Rate, Gender, and Peak-to-Peak Segment.

The data structure diagram (Figure 48 in Section 4.8.3) shows that the random effects for the timing error data have a three-way structure: (a) peak-to-peak segments are nested in their respective samples, (b) samples are nested within experimental conditions, and (c) samples from all experimental conditions are nested in the corresponding participants. Section 4.3.2.2, *Expanding the Random Intercept Linear Mixed Model to a Three-Way Data Set,* discussed previously, described a three-way hierarchical model structure as Equation 4.4:

$$Y_{ijk} = \alpha + a_i + b_{j|i} + \varepsilon_{ijk} \qquad \text{(Equation 4.4)}$$

Equation 4.4 means that this model includes the following components:

(a) overall average for the RMSE values represented by α;

(b) average deviations from the overall RMSE values for each participant represented by a_i;

(c) an average deviation for each trial (nested in the corresponding participant) from the average for the corresponding participant represented by $b_{j|i}$; and

(d) the error terms, denoted by the term ε_{ijk}, which reflect individual observations that deviate from the fitted RMSE values at each peak-to-peak segment.

Note that no details are included in this model about the experimental conditions and explanatory factors.

Applying Equation 4.4 to the present data set, the term a_i is an intercept specific for each participant (i_1-i_{35}), and the $b_{j|i}$ is an intercept for each trial $_j$ (j_1-j_{16}), nested in the corresponding participants. Both terms are assumed to be normally distributed with mean 0 and variance σ_a^2 and σ_b^2, respectively. ε_{ijk} is the residual error after accounting for participant and trial effects and is specific to each observation (ijk_1 - ijk_n), reflecting the deviations from the fitted values at each peak-to-peak segment. ε_{ijk} is assumed to be normally distributed with mean 0 and variance σ^2. Note that this model specifies the hierarchical relationships among the observations but does not specify the experimental conditions and explanatory factors in this analysis.

As with the wiggle and timing error data, the first linear mixed effects model for RMSE values was based on the graphical displays of the data and the underlying questions of the research regarding: (a) the effect of concurrent visual feedback (Feedback:Absent and Feedback:Present conditions) on spatial errors; (b) the robustness of those effects with a timing perturbation (i.e., the TargetRate:Changing condition); (c) trends in those effects over time, shown by changing effects over the successive peak-to-peak segments in the sample; and (d) differences in performance between genders.

Thus, the initial RMSE model contained these factors:

(1) Feedback, with values absent and present;

(2) Target Rate, with values changing and constant;

(3) Gender, with values female and male;

(4) Peak-to-peak segments in the sample (peaks 1-2, peaks 2-3,…, peaks 11-12)

The model also has to allow for the possibility that the explanatory variables may interact, meaning that the effect of any explanatory variable (Feedback, Target Rate, Gender, or Peak-to-Peak segment) on RMSE values may depend on the level of another explanatory variable. Finally, RMSE values appear to change nonlinearly in the Feedback:Absent-TargetRate:Changing condition in Figure 57, suggesting that the nonlinearity in the analysis of spatial errors may require a more complex model than a simple linear mixed effects model.

4.9.1. Step 1: Evaluating the Need for a Linear Mixed Model

After evaluating exploratory plots and building an initial model structure with the four-way interaction Feedback∗TargetRate∗Gender∗Peak-to-Peak Segment to assess the data, the next step requires a comparison of a fixed effects model and the linear mixed effects model. This step allows us to determine whether a mixed effects model is necessary for proper specification of these data, or if a fixed effects model would suffice (West et al., 2007; Zuur et al. 2009). Since the only difference between the fixed model and the linear mixed effects model is the random component, these models are nested and can be compared by calculating the likelihood ratio (Zuur et al., 2009) using ANOVA.

The fixed model for spatial errors is denoted by Equation 25, *RMSE-gls. RMSE-lme* represents the linear mixed effects model with the same four-way interaction as fixed effects, and random effects consisting of participants, denoted by $a_{i,}$ with trials nested in participants, denoted by $b_{j|i,}$ in Equation 26.

RMSE-gls: $RMSE_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + \varepsilon_{ijk}$ (Equation 25)

vs.

$$RMSE_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + a_i + b_{j|i} + \varepsilon_{ijk} \qquad \text{(Equation 26)}$$

The comparison of the fixed effects model (Equation 25) and linear mixed effects model (Equation 26) is displayed in Table 4. The fit of these models can be compared by calculating the likelihood ratio, denoted by *L. ratio,* using ANOVA (Zuur et al., 2009).

Table 4

Comparison of RMSE Data Using Generalized Least Squares and Linear Mixed Effects Models

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
RMSE-gls	1	89	-3926.32	-3332.44	2052.16			
RMSE-lme	2	91	-7018.87	-6411.65	3600.44	1 vs 2	3096.55	<.0001

The likelihood ratio statistic (L = 3096.55, p <.0001) shows that the two models are significantly different. The reduced *AIC* and *BIC* values for the linear mixed model indicate that the linear mixed model with random effects provides a significantly better fit to the RMSE data than the fixed model structure.

4.9.2. Step 2: Initial Modeling Procedures: Developing the Beyond Optimal Model for Spatial Error Data

After determining that the linear mixed model was more appropriate than the fixed effects model for the RMSE data, I developed an initial model based on the research questions, the data structure, and the exploratory plots as discussed in Sections 4.8.3 and 4.8.4. Of 5984 experimental units, eight outliers were removed from the data set used for analysis.

The first step in model development involves setting the structure for the fixed effects. Instead of using mathematical notation, I am writing the fixed portion of the model in words, after Galwey, (2006), as represented by Formula 18

$$\text{Feedback} * \text{Target Rate} * \text{Gender} * \text{Peak-to-Peak Segment} + \text{Trial}_{\text{condition}} + \text{Trial}_{\text{session}} \quad \text{(Formula 18)}$$

The series of * symbols in this model indicate a four-way interaction. The four-way interaction in Formula 18 accounts for changes in each explanatory variable at each level of the other explanatory variables. Since terms enter the model hierarchically, a mathematical rule requires all main terms, two-way interactions, three-way interactions, and four-way interactions are included in this model.

The random portion of the linear effects model must include an intercept for participants (a_i) to allow us to subject-specific effects for participants and a random intercept for trials ($b_{j|i}$) for estimating trial-specific effects, nested in the corresponding participants (intrasubject variability).

Combining the fixed and random parts of the model and applying it to the factors in this study, we have the following model in Equation 27:

$RMSE_{ijk}$ = Feedback_{ijk} * Target Rate_{ijk} *$\text{Peak-to-PeakSegment}_{ijk}$ * Gender_{iik} +

$\text{Trial}_{ab} + a_i + b_{j|l} + \varepsilon_{ijk}$ (Equation 27)

The linear mixed effects model indicates that $RMSE_{ijk}$ refers to each RMSE value at each Peak-to-Peak Segment; and Feedback_{ijk}, Target Rate_{ijk}, Gender_{ijk}, and $\text{Peak-to-PeakSegment}_{ijk}$ refer to the effect of Feedback, Target Rate, Gender, and Peak-to-PeakSegment (2-12) on RMSE values on each response, nested in each participant. Trial_{ab} is evaluated separately to determine if trends occurred over the trials in each experimental condition, *a,* (1-4) **or** over the experimental session*, b,* (1-16)**.** Participants are entered as random effects. Thus the analysis of this model will assess spatial copying errors at each peak-to-peak segment of the samples, and evaluate the interaction of Feedback, Target Rate, and Peak-to-Peak segment on the RMSEs, allowing for separate effects for females and males. The random components are reflected in the terms, a_i and $b_{j|l}$ which allow a separate intercept for participants (a_i) and for trials nested in participants ($b_{j|l}$). The residuals, ε_{ijk}, are obtained from the differences between the *k*th observation and the corresponding fitted values for the RMSEs within each trial, nested in the corresponding participant.

4.9.2.1. Exclusion of Speed as a Predictor

Sample speeds were excluded as predictors because of the presence of endogeneity, as discussed previously in the sections on wiggle and timing errors. Sample speed can be controlled to a limited extent by the experimenter by the pre-set speed of the target, but problems arise when participants observe their spatial errors when copying the target shapes. In the present study, they manipulated their drawing speeds to improve either spatial or timing accuracy when concurrent visual feedback was present. As a result, sample speeds can function as both a predictor and an outcome of participant responses, producing endogeneity, which cannot be adequately addressed in this model framework. Endogeneity was least likely to occur when concurrent visual feedback was not provided and the target rate was constant, so sample speeds were observed in that condition to assess whether the results were consistent with Fitts' Law (Fitts, 1954), which would provide a measure of external validation for the software operation.

4.9.2.2. Initial Modeling Procedures: Checking For Order Effects in Spatial Error Data

To begin the model development process, I assessed the effects of trial order to determine whether the data depended on the order of the four trials within each experimental condition, *a,* (1-4) or the sixteen randomized trials in the experimental session, *b,* (1-16). Thus the first model that was analyzed included the four-way interaction among Feedback, Target Rate, Gender, and Peak-to-Peak Segment, with trial order within each experimental condition and trial sequence over the experimental session added separately as fixed effects (Equation 4.23). Participants, with trials nested in participants, comprised the random effects in this model, as indicated by the data structure diagram. This model was shown previously by Equation 27.

Trial order within each experimental condition, *a* (1-4), was not found to be significant statistically (F= .2127, p = .8876) in the linear mixed model defined by Equation 27. Trial sequence over the set of trials over the experimental session, *b* (1-16) was nonsignificant statistically as well (F= .6499, p = .8333). Since trial orders in each experimental condition and across the experimental session, were not found to be significant, the terms of the *beyond-optimal* linear mixed effects model (Zuur et al., 2007) included the four-way interaction between Feedback, Target Rate, Gender, and Peak-to-Peak Segment. Of the various possibilities for the choice of random effects, participants and trials nested in participants yielded the lowest AIC. Thus, participants comprised the random effects in the model with trials nested in the corresponding

Participant. This framework enabled an assessment of intrasubject variability across trials and within trials. Thus, the first model tested, the *beyond-optimal* model was the four-way interaction with the nested random terms of trial nested in participant, as shown by Equation 28.

$$RMSE_{ijk} = Feedback_{ijk} * TargetRate_{ijk} * Peak\text{-}to\text{-}Peak\ Segment_{ijk} * Gender_{ijk}$$
$$+ a_i + b_{j|i} + \varepsilon_{ijk}$$

(Equation 28)

4.9.3. Step 3: Initial Modeling Procedures: Checking for Constant Variance

Once the initial model has been chosen, the next step in model development involves plotting and assessing the model residuals for constant variance. In the spatial error data, the residuals from the full model denoted by Equation 28 displayed structure that was not fully accounted by a linear mixed model framework, as shown in Figure 58.

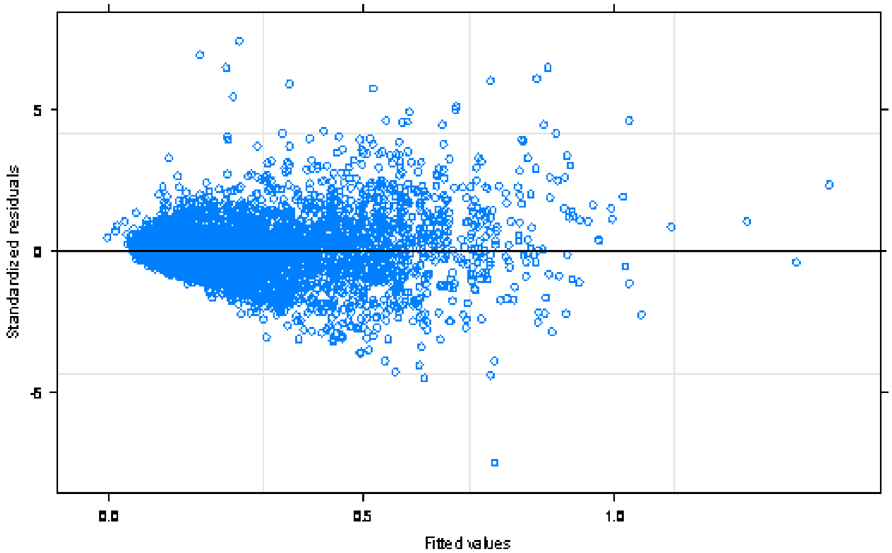

Figure 58. Residuals from the *beyond-optimal* linear mixed effects model for RMSE values.

The residuals in Figure 58 have a widening wedge pattern, indicating that the values of residuals broaden as fitted values increase. Further, since standardized residuals have variance = 1, many of the residuals in Figure 58 are quite large. Thus, even though the linear mixed model was an improvement over the fixed model, the residual pattern in Figure 58 indicates that a linear mixed model does not adequately resolve the dependency in the spatial error data.

In such situations of non-constant variance, several possibilities remain available for reducing heteroscedasticity in the residuals: transformations, adding exponentials, and creating interactions (Oehlert, 2000; Zuur, 2009). A log transformation of the response resulted in improved variance, as shown in Figure 59.

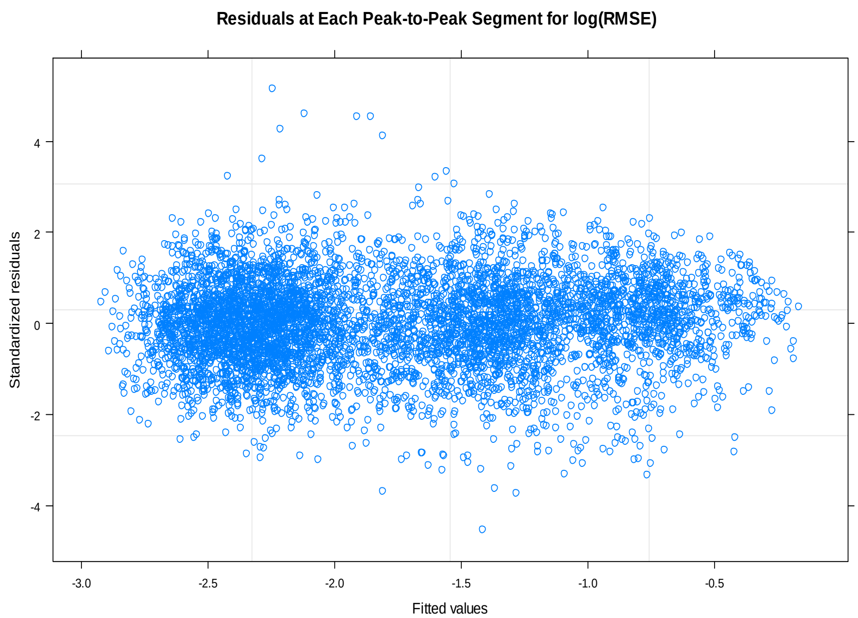

Figure 59. The residual plot of the same data set as in Figure 58, using the log of RMSE as the response.

In addition to these options, mixed model structures permit adjustments to the covariance structures to be adjusted, random slopes to be modeled, or smoothers to be applied, creating a Generalized Additive Mixed Model (GAMM). In the present case, all possible interactions are already included in the model, and further experimentation showed little improvement in the residual patterns by adding exponential terms. Using the log of the response, log (RMSE) significantly reduced the non-constant variance. Exploring alternate structures for the covariances, modelling random slopes, or applying smoothers in a GAMM are appropriate next steps to find the optimal model.

4.9.3.1. Selecting a Covariance Structure

The method of adjusting covariance structures involves lifting the assumption that all residuals ε_{ijk} are normally distributed with mean 0 and variance σ^2. An advantage of mixed models over fixed models is that linear mixed models can allow for heterogeneity by using multiple variances σ^2. For example, if males have more variation in spatial errors than females, we can use two variances, σ_{male}^2 and σ_{female}^2. Similarly, we can use different variances for different levels of Feedback (absent and present), Target Rate (changing and constant), and for different peak-to-peak segments across the sample.

Using different variances also allows us to apply different types of variance structures for the correlations among the data. According to West et al. (2007), the choice which variance structure to choose requires consideration of at least two factors:

(1) Residual box-and-whisker plots show that the variation differs with the levels of Feedback, Target Rate, Gender, and peak-to-peak segments.

(2) Different variance structures can be compared for better fit, using REML estimation and AICs.

Both options (1 and 2, above) may involve considerable experimentation to determine the best fitting model (Zuur, Ieno, & Elphick, 2009). Since Zuur et al. (2009) advises the varIdent structure when predictors are categorical, as they are in the present data set, the varIdent structure was assessed with these data. The varIdent structure imposed on feedback improved the fit over previous models, but was only partially successful in resolving the non-constant variance, such that the reliability of the statistical tests was still in question.

4.9.3.2. Assessing the Need for a Generalized Additive Mixed Model

Following the addition of the varIdent covariance structure, I examined the residuals for the spatial data to determine if variations in residuals were related to patterns in any of the exploratory variables of Feedback, Target Rate, the Peak-to-Peak Segments (1-2, …, 11-12), and Gender, and found that the greatest variability occurred in the TargetRate:Changing condition, where data had a highly nonlinear relationship with the outcome variable, RMSE, as shown in Figure 58. Nonlinearity of this type may be handled with a Generalized Additive Mixed Model.

Smoothing function for peak - changing Rate

Feedback absent (top) and present (bottom)

Smoothing function for peak - constant Rate

Feedback absent (top) and present (bottom)

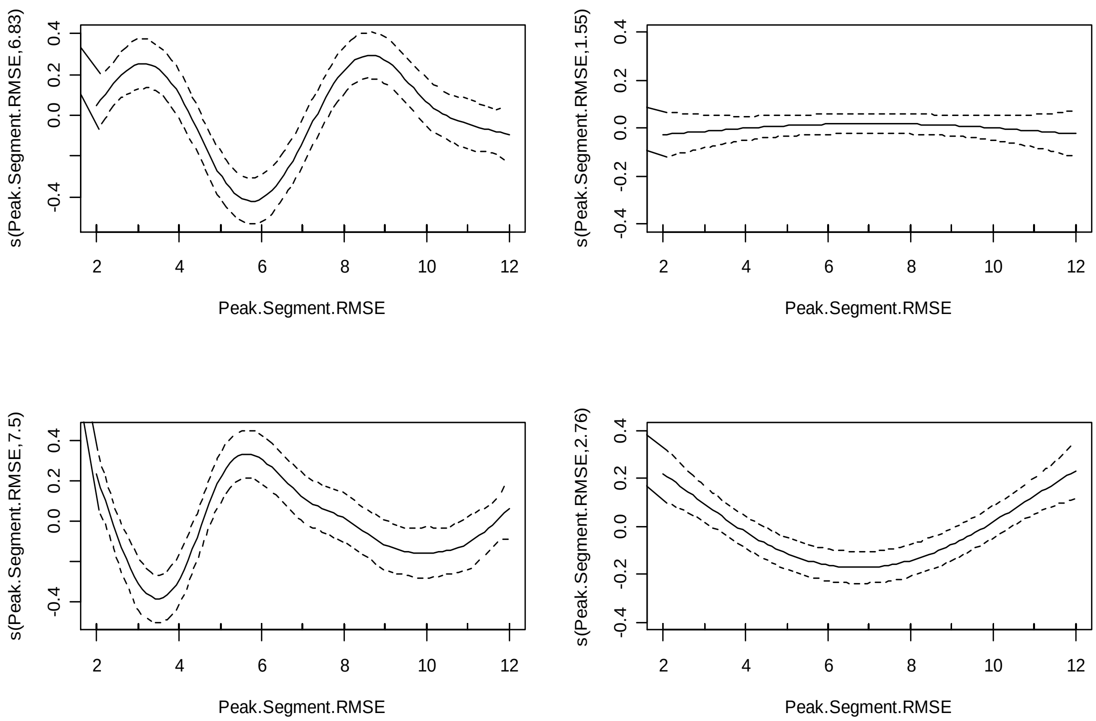

Figure 60. The effect of smoothed peak segment term, *s*(peak.Segment.RMSE) on RMSE. The left panels represent Target Rate:Changing conditions, TargetRate:Constant conditions are on the right. The upper panels (upper left and right) correspond to RMSEs from Feedback:Absent trials, while the lower panels (lower right and left) correspond to RMSEs from the Feedback:Present conditions. These plots are taken on the non-transformed data, not the logarithm of the response. Using a log transform yields plots that are very similar, although the estimated degrees of freedom differ.

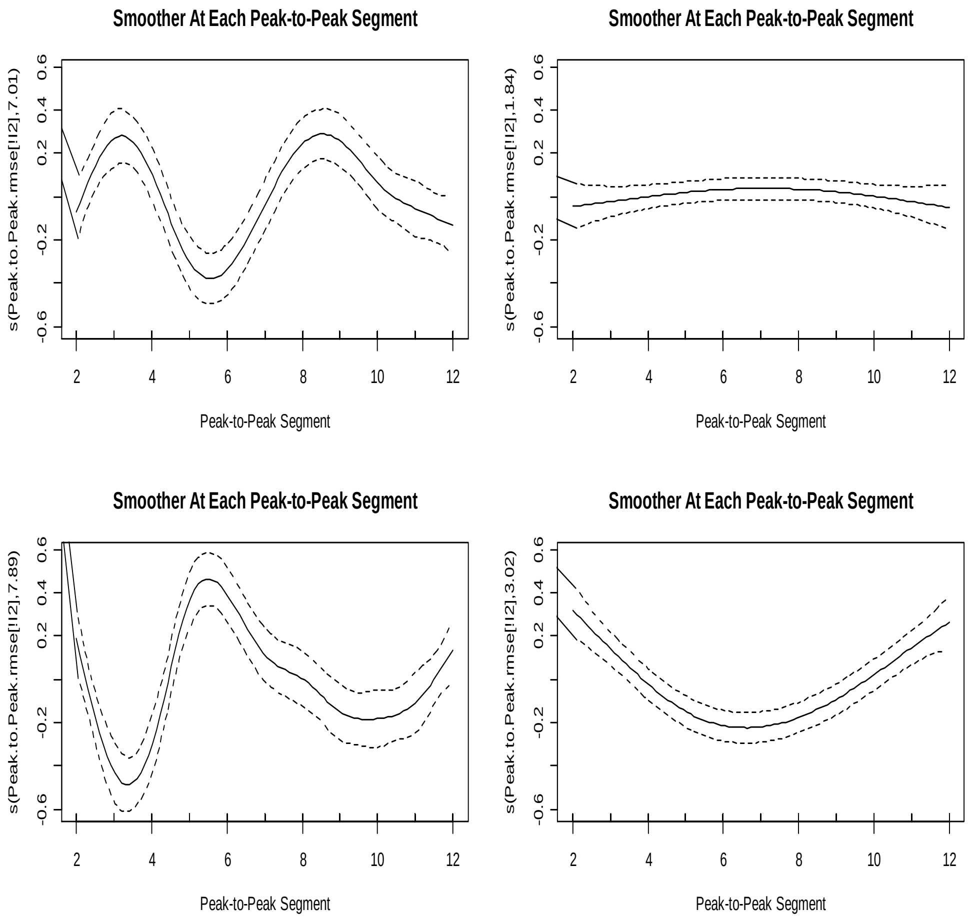

Figure 61. Smoothers applied to the log of the spatial error data, categorized by experimental treatment.

Dealing with nonconstant variance: smoothing methods. The addition of a smoothing function at each peak-to-peak segment in each experimental condition was useful for determining the cause of the wedge-shaped residual pattern from the original model. As displayed in the left panels of Figure 61, the TargetRate:Changing condition showed the greatest variability across peak-to-peak segments for both Feedback:Absent and Feedback:Present conditions. Less variability is observed in the Feedback:Present-TargetRate:Constant condition (lower right panel of Figure 61), where the smoothing function increases at the ends of the samples, suggesting that greater spatial errors occurred at the beginning and end of the samples in that condition. The lowest variability in RMSE values was found in the Feedback:Absent-TargetRate:Constant condition, as indicated by the relatively straight line the top right panel of Figure 61, suggesting that these data were nearly linear.

The values on the y-axes in Figure 61 show the cost of smoothing by indicating the estimated degrees of freedom. The higher the estimated degrees of freedom, the more nonlinear the data represented in the plot. The values on the y-axis for the Feedback:Absent-TargetRate:Changing condition (top left panel of Figure 61) show that the effect of the Peak-to-Peak Segment was significant on 6.83 estimated degrees of freedom ($p < .0001$). As shown in this plot, the RMSE values varied greatly at different peak-to-peak locations on the sample. The smoothed Peak-to-Peak Segment term was also significant in the Feedback:Present-TargetRate:Changing condition (lower left panel in Figure 61) on 7.5 estimated degrees of freedom ($p < .0001$). These estimated degrees of freedom and their associated p-values indicate that the RMSE values in those panels were significantly non-linear, varying considerably by the location of the Peak-to-Peak Segment. The nonlinearity is most noticeable from approximately Peak-to-Peak

Segment 3-8 in the TargetRate:Changing conditions, which correspond to the abruptly changing target rate in that section of the sample.

The effect of the smoothed peak Segment term, *s*(peak.Segment.RMS) in the Feedback:Present-TargetRate:Constant condition on 2.76 estimated degrees of freedom is highly significant ($p < .0001$). The shape of the plot shows that variability in the RMSE values occurred at the Peak-to-Peak Segments at the beginning and end of the sample, although the RMSE values were consistent in the middle of the sample. The effect of the smoothed Peak-to-Peak Segment term in the Feedback:Absent-TargetRate:Constant condition on 1.56 degrees of freedom was not significant ($p = .7$) indicating that the RMSE values in the Feedback:Absent-TargetRate:Constant condition were consistent (i.e., approximately linear) across all the Peak-to-Peak Segments in the sample.

4.9.4. Step 4: Initial Steps in Building a Generalized Additive Mixed Model for the RMSE Data, using log(RMSE) as the response

4.9.4.1. Selecting a Structure for Fixed Factors

The results of the data exploration thus far indicate that a mixed model is necessary to model the participant effects, using random intercepts for participants and a generalized additive mixed effects structure to deal with the nonlinearity in the RMSE data . A variance stabilizing mechanism, termed the *varIdent* correlation structure, was used. This structure allows the researcher to relax the constraints on categorical variables, so that each treatment (i.e., each feedback-target rate combination) has its own variance. In the present case, different visual conditions produced considerably different variances: where participants made fewer spatial errors when they could see their own response to the copying task, the variability in performance was much broader when the participants could not observe their own efforts at copying the target motion. Thus each level of feedback (absent, present) was allowed separate variances in the model. Following the selection of a covariance structure, the next step in model development is to build a generalized additive mixed model (GAMM) as described in Section 4.3.

Since the difference between the linear mixed effects model and the GAMM is the introduction of smoothers to improve the fit to the data, the fixed factors remain the same as in the previously tested linear mixed effects model:

Feedback * Target Rate * Gender * Peak-to-Peak Segment (Formula 18)

Figure 61 indicated that Feedback (Absent and Present) and Target Rate (Changing and Constant) produced different responses at different peak segments. Previous exploratory plots indicated that the predictors of Feedback, Target Rate, and Gender might have had interactive effects on the responses, so the GAMM must also include interactions in the fixed model. In addition to these fixed effects, the GAMM must include a smoother (f_{ftg}) to account for the changing variation of Feedback, Target Rate, and Gender at each Peak-to-Peak segment. This is done by adding indices for Feedback, Target Rate, and Gender to the smoother at the Peak-to-Peak Segments. Note that the model depicted by Equation 29 contains fixed effects only, and does not include a random component.

$$log(RMSE_{ijk}) = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + f_{ftg}(Peak) \qquad \text{(Equation 29)}$$

The indices *ijk* denote the *k*th observation in the *j*th trial ($j_{1\text{-}16}$) which is nested in the corresponding *i*th participant. The *k*th observation is obtained at each peak-to-peak segment in the sample. The index f_{ftg} *(peak)* stands for the smoother applied to peak-to-peak segment, allowing it to have different values based upon the values for Feedback (*f*), Target Rate (*t*), and Gender (g). Smoothers in this model, indicated by f_{ftg} before (*Peak)*, are not a parametric term in a model, but are depicted as a line in the data plot that represents an optimized participant response at each peak, allowing each Feedback-Target Rate-Gender combination to have its own peak effect on log(RMSE) values.

4.9.4.2. Selecting a Structure for Random Effects

Once the fixed effects are determined, the next step in model development involves discovering the most appropriate structure to handle the correlated data. Typically, this process requires selecting random effects from the repeated observations and determining a structure for them, which may require experimenting with random intercepts, random slopes, and covariance structures. As described previously, competing models were compared using *AIC* values to assess which structures fit the data best. Using reduced *AIC* values for the decision rule, the best fitting GAMM for the RMSE data involved nesting trials in participants and adding a varIdent structure. Due to convergence problems, the varIdent structure included only Feedback (Absent and Present), which appeared to have the greatest influence on the outcome variable.

After experimenting with different types of random factor arrangements once the smoother was applied, the best fit was obtained when participants were permitted to have different intercepts (i.e., random intercepts), as described previously by Formula 14.

$$a_i + b_{j|i} + \varepsilon_{ijk} \qquad \text{(Formula 14)}$$

Adding random slopes did not provide any significant improvements to the fit of this model. Instead, the problem of non-constant variance with these data appeared to be adequately resolved by the application of smoothers in the GAMM, the varIdent correlation structures, and random intercepts. Summarizing, the non-linearity and non-constant variance in the RMSE data were addressed after considerable experimentation by the following approaches:

(1) Random effects (intercepts) were provided for participants, with trials nested in participants.

(2) Smoothers in a GAMM format reduced the nonlinearity in RMSE values

(3) The VarIdent correlation structure allowed Feedback to have separate variances at each level (Absent or Present), and further reduced nonlinearity in the model.

These approaches were specified with subroutines from the *nlme* package (Pinheiro et al., 2008) and the *mgcv* package (Wood, 2004, 2006) in *R.* Thus, the *beyond-optimal* GAMM for the RMSE data included a four-way interaction among Feedback, TargetRate, Gender, and Peak-to-Peak Segment; a smoothed term for Feedback at each Peak-to-Peak segment; random effects comprised of participants with trials nested in participants and participant-specific intercepts; and a varIdent covariance structure. This model is depicted by Equation 30.

$$\log(RMSE_{ijk}) = Feedback_{ijk} * TargetRate_{ijk} * Gender_{ijk} + f_{ftg}(Peak) + a_i + b_{j|i} + \varepsilon_{ijk},$$
$$\text{with } \textit{varIdent} \text{ structure for Feedback}$$

(Equation 30)

This model includes participant-specific intercepts and trial-specific intercepts for each participant. Note: Since the peak term is smoothed, it cannot be part of the interaction among the other predictors.

The indices in Equation 30 mean that the random effect terms, a_i and $b_{j|I}$ allow for a separate intercept for participants (a_i) and for trials nested in participants ($b_{j|I}$). The residuals, ε_{ijk}, are obtained from the differences between the *k*th observation and the corresponding fitted values for log(RMSE) observations in each trial.

4.9.4.3. Evaluate the Initial Beyond-Optimal GAMM for Constant Variance

The goal of this step is to determine whether the generalized additive mixed model strategy reduced the heteroscedasticity that was evident in the linear mixed model depicted in Figure 61. The *beyond-optimal* GAMM denoted by Equation 30 produced the residual pattern shown in Figure 62.

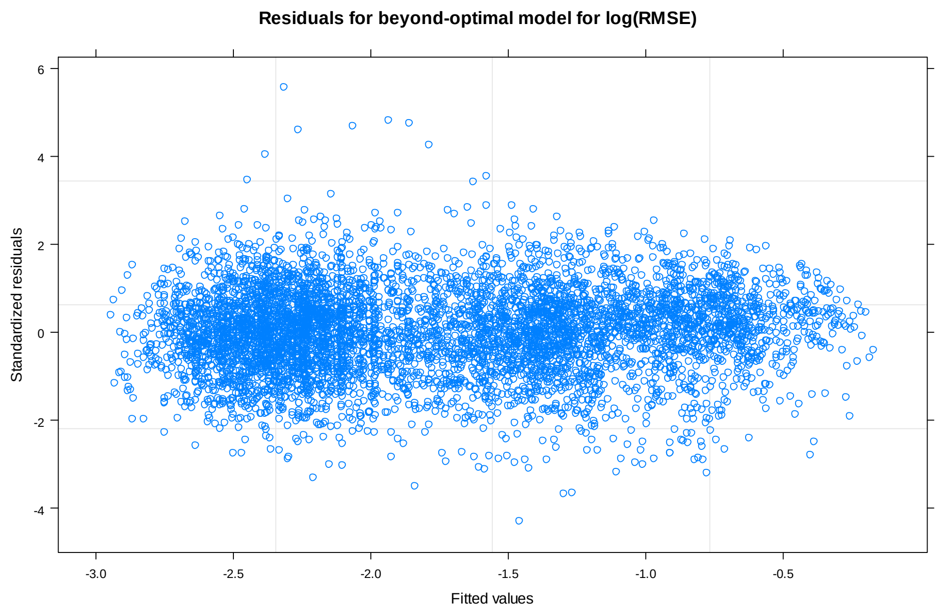

Figure 62. Initial *beyond-optimal* model of log(RMSE) data, using a generalized additive mixed model framework, with a varIdent covariance structure for feedback and participants with trials nested in participants as random effects.

In Figure 62, the residuals for the log of the response (RMSE) appear to be randomly distributed across zero with a reasonably random structure (Zuur, personal communication, March 18, 2009), although some outliers remain, which were associated with one participant. However, the reasonableness of the pattern allows us to assume that the generalized additive mixed model denoted by Equation 30 is a suitable starting point for analyzing these data.

4.9.4.4. Determining the Effect of Trial Order over the Experimental Session

Although trial order was tested for the initial linear mixed effect model for the RMSE data, it seemed prudent to plot the residuals of the GAMM against the trial order (trial 1-16) and to test trial order as a fixed effect to determine any order effects with the GAMM model framework, as shown in Figure 62. With the exception of some outliers, the values for the residuals of the *beyond-optimal* GAMM were concentrated around zero for all sixteen trials, indicating that trial order did not appear to have a significant effect on the RMSEs. The order of trials was not found to be significant across experimental conditions, *a,* (1- 4) within experimental conditions (F = 0.1237, p = 0.946), or across the experimental session, *b,* (1-16) (F = 0.656, p = 0.418). These findings do not support including include trial orders as fixed effects in the beyond-optimum GAMM for the RMSE data, but the structure of the data indicated that trials should be modeled as a random effect nested in participants. Thus, the constant variance and insignificant effects of trial orders support the decision to use the model in Equation 30 as the initial GAMM for fitting the spatial error data.

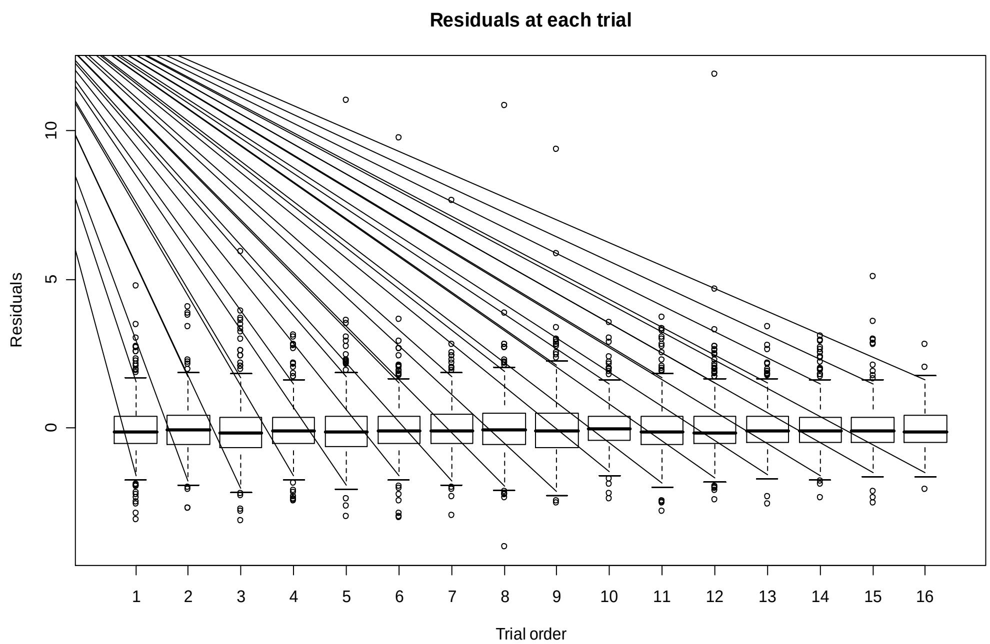

Figure 63. Residuals at each trial in sequence over the experimental sessions for the *beyond-optimal* GAMM for RMSE values.

4.9.5. Step 5: Refining the model for spatial accuracy data

Following the model selection procedures outlined in West et al. (2007) and Zuur et al. (2009) which use the Akaike Information Criterion (*AIC*) as the standard, I sequentially dropped statistically nonsignificant terms from the *beyond-optimal* GAMM based on *p*-values, using stepwise backward selection. Beginning with the GAMM that included Feedback, Target Rate, Gender, the interactions among these predictors (Feedback∗TargetRate∗Gender), and the smoothed Peak-to-Peak Segment term

($f_{ftg}(Peak)$), i.e., Equation 30, model simplification proceeded by removing the following terms in order with refitting at each step:

(1) The three-way interaction term, Feedback∗Target Rate∗Gender, was found to be non-significant (F = 0.003, p = 0.9581).

(2) Two two-way interaction terms were found to be non-significant:

(a) Feedback:Gender (F = 0.451; p = 0.5019) and

(b) Target Rate:Gender (F = 0.557, p = 0.4554).

(3) The main effect Gender was found to be non-significant (F = . 546,

p = 0. 4602). Since Gender was no longer significant as a main effect, it was removed from the smoother for peak as well. Tests on the smoother confirmed this choice.

Note that the terms that were removed were consistent with Figure 61, the exploratory plot for the three-way interaction among Feedback∗TargetRate∗Gender. Thus, the final optimal model was given by Equation 31, using the log of RMSE as the response:

$$Log\,(RMSE_{ijk}) = Feedback_{ijk} + TargetRate_{ijk} + Feedback_{ijk} * TargetRate_{ijk} + f_{ft}(Peak) + a_i + b_{j|i} + \varepsilon_{ijk}$$

Equation 31

The model depicted by Equation 31 allows different variances for different levels of Feedback, a smoothed peak term for feedback and target rate, random effects comprised of participants and trials nested in participants, allowing for participant-specific and trial-specific intercepts.

Since the model in Equation 31 can be difficult to understand in practical terms, I am presenting the syntax for this model in the statistical language *R* in Formula 19.

```
Log(RMSE) <- gamm(RMSE ~ Feedback*Target Rate
+ s(peak.Segment, by = as.numeric(Target Rate =="changing" & Feedback =="absent"))
+ s(peak.Segment, by = as.numeric(Target Rate =="constant" & Feedback =="absent"))
+ s(peak.Segment, by = as.numeric(Target Rate =="constant" & Feedback == "present")),
+  random= list(participants =~1, trials= ~1),  weights=varIdent(form=~1|Feedback)),
method = "REML", na.action = na.omit)                              (Formula 19)
```

This syntax indicates that the final optimal model for RMSE contains Feedback and Target Rate and the interaction between them (Feedback∗Target Rate), with smoothed peak terms, allowing peak-to-peak segments to have different values, depending on the levels of Feedback (Absent or Present) and Target Rate (Changing or Constant). As a result, the GAMM includes three smoothers. The smoother for Feedback:Present-TargetRate:Changing was removed from the model because it was not significant (*p* =

The term *random= list(participants =~1, trials= ~1)* indicates that participants are entered as random effects, and that the levels of the trials (1-16 across the experimental session) are nested in participants, meaning that Formula 19 represents a random intercept model. Thus, the random structure in the GAMM accounts for correlations at two levels: (a) between observations at different peak-to-peak segments within the same trial and (b) between trials from the same participant.

A *varIdent* structure was imposed on Feedback, allowing for different variances for Feedback:Absent and Feedback:Present conditions. The model was fit using restricted maximum likelihood (REML).

4.9.6. Step 6: Interpret Significance of the Smoothers in the RMSE Model

Table 5 displays the significance of the smoothing terms used in the GAMM for the spatial error data. The abbreviation *edf* in Table 5 denotes the estimated degrees of freedom, and is printed in Figure 61 as the calibration value for smoothing on the y-axis in the plots, meaning the estimated degrees of freedom (*edf*) required to fit the data. The values of the smoothers in these data range from 1, representing a straight line, meaning that the data are consistent (i.e., linear) at each measurement, and a value of 6.924 which means that the smoother requires more degrees of freedom to fit data that are considerably non-linear. Table 5 also displays the value of the *F*-test for the smoothers, which are used to determine whether the value of the statistical test for the smoother is different from 0, as indicated by p-values. These tests assess the necessity of the smoothers to for modeling the data by comparing to the null hypothesis, smoothers = 0.

Table 5

*Approximate Significance of Feedback*TargetRate Smooth Terms in Optimal Model for Spatial Accuracy Data (log RMSE)*

Smoother	edf	F	p-value
s(peak):Smoother: absent, changing	1.001	6.822	0.0011
s(peak):Smoother: absent, constant	1.000	8.681	0.0002
s(peak):Smoother: present, constant	3.576	10.162	<0.0001

Note that the s(peak):Smoother: present, changing is not included in Table 5 because it was found to be non-significant (p = 0.0968)

The results in Table 5 demonstrate that the Feedback:Target Rate smoothers were found to be statistically significant, suggesting that each specific experimental condition contributes to the non-linear pattern. However, Zuur (2009) cautions that the estimated edf values for smoothers are approximate, particularly for with p-values close to 0.05. These issues were being addressed by planned updates to the mgcv package (Wood, 2009).

These results also show that the data in the Feedback:Present-TargetRate:Changing condition were effectively linearized with the optimal GAMM model described by Equation 31 and Formula 19. Nonlinearity in this model was most evident in the Feedback:Absent-TargetRate:Changing condition. Further, the significant *p*-values indicate that each combination of predictors (Feedback:Target Rate) is necessary for an adequate representation of these data (Crawley, 2007).

4.9.7. Step 7: Final Tests to Validate log(RMSE) Model

4.9.7.1. Assessment of Residuals of Final GAMM for log(RMSE) Values

After the RMSE data were fit with the optimal GAMM random intercept model with varIdent structure for feedback shown by Equation 31, the residual plot for this model, shown in Figure 63, was assessed for constant variance.

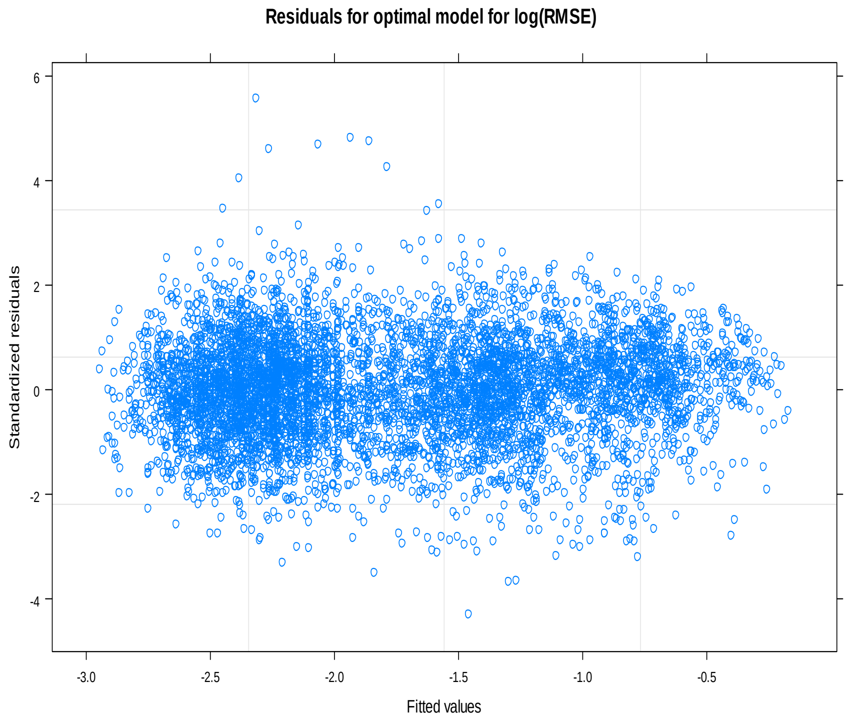

Figure 64. Residuals of optimal random-intercept GAMM for the log-transformed RMSE data, with varIdent structure for Feedback, as specified by Equation 31.

Despite a few outliers, the residuals appear evenly spaced over the zero-value across the plot, indicating that a reasonably random distribution of residuals was achieved with the final model, and suggesting that the statistical tests with this model are reliable. The histogram of the residuals confirms the presence of outliers with a right skew, but is otherwise reasonably balanced. The plots of residuals in Figure 64 and Figure 65 below indicate that the optimal GAMM model in Equation 31 accounts for much of the structure in the data and is a reasonable fit, although some poorly fit data remain.

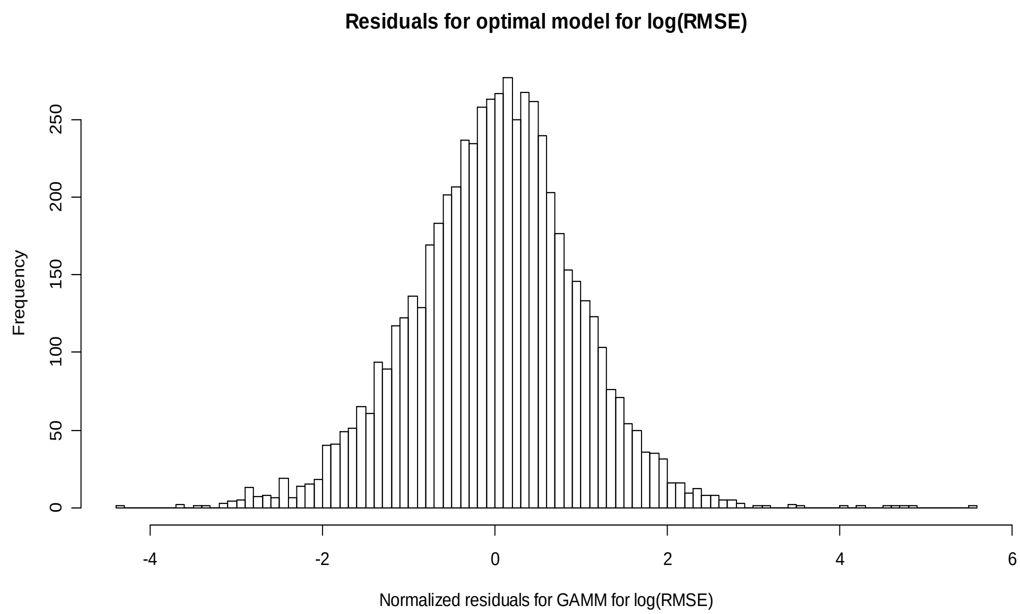

Figure 65. Residuals from the optimal model for the RMSE data in Equation 31.

Figure 65 shows residuals for the final model for the spatial error data, categorized by experimental conditions and peak-to-peak segment. The residuals in these four plots show few unusual features, which supports the hypothesis that the final reduced model is an adequate summary of the log(RMSE) values in the current data set. The extreme values in this residual plot suggest that further data exploration may be important to determine if these residuals share any particular characteristics. Since these data were contributed by a very few participants who responded consistently quite differently from the rest of the participants, they were retained in the data set. Had these results become apparent when the participants were still available, a discussion with these participants may have been useful to determine what may have contributed to the extreme responses they produced.

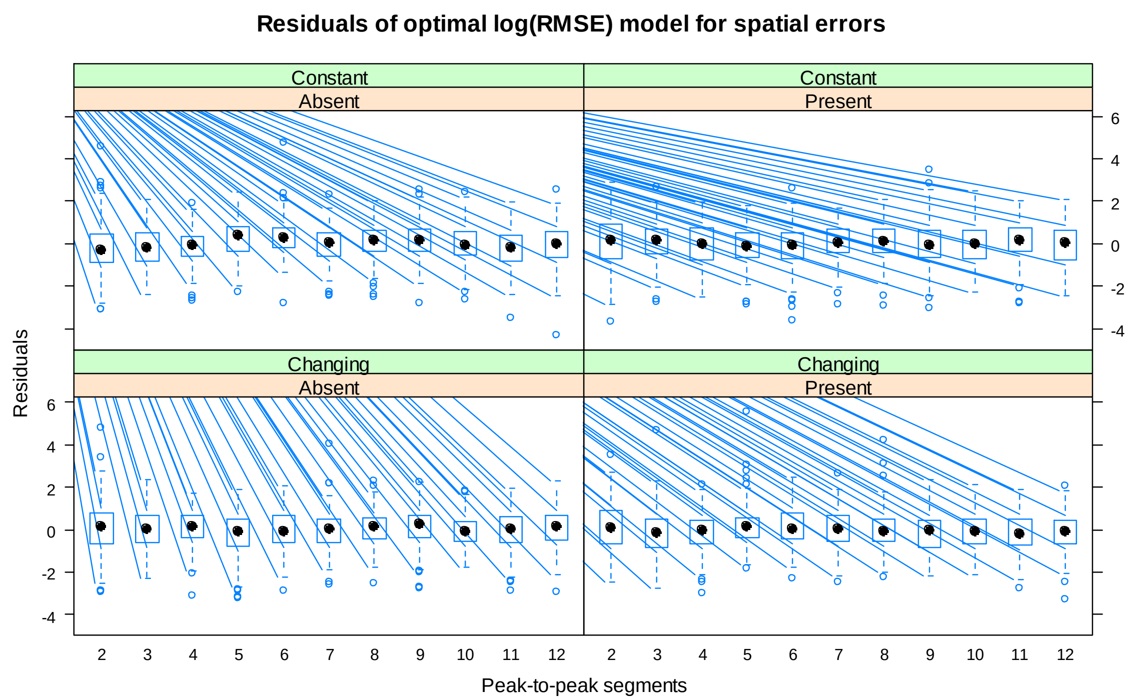

Figure 66. Residuals for optimal log(RMSE) model, categorized by experimental condition in the four quadrants which divide the data into peak-to-peak segments within the sample, displayed on the *x*-axis.

4.9.7.2. Assessment of Akaike Information Constants for Alternate RMSE Models

In addition to the results of the residual plots, the Akaike Information constant (*AIC*) was compared in the beyond-optimal and final reduced GAMM for the spatial error data. The *AIC* for the final reduced GAMM for the log(RMSE) responses (-9810.517) was lower than the *AIC* for the full, beyond-optimal model (-9804.806), indicating that the reduced GAMM is superior to the beyond-optimal model for these data.

4.9.7.3. Interactions

In Section 4.5.5.2, I stated that the interpretation of main effects involved in interactions are often not possible, and, at best, unreliable (Underwood, 1997) except in very rare cases, such as when one effect considerably dwarfs the other (Oehlert, personal communication, June 5, 2009). Deletion tests allow for estimation of main effects by using likelihood ratio tests to determine whether the main effects have statistically significant effects on the response. The results of the deletion tests are presented in *Chapter 5*, *Results*; they provide evidence that both Feedback and Target Rate are necessary in the model. Thus, even though we cannot interpret the variables Feedback and Target Rate directly as main effects due to the interaction in the final model, this process of comparing nested models with a deletion test provides approximate information about the significance of predictor variables when they are included in an interaction.

4.9.7.4 Intraclass Correlations within Spatial Error Data

As described in Section 4.2.1, random effects in the present study introduce correlations between observations from the same participant. In the present study, intraclass correlations for participants and trials nested in participants were calculated to assess the similarities within those subgroups, which reflected the hierarchical data structure shown previously in Figure 48. The calculated proportion of variance in the unit of analysis (i.e., within-subjects variability and within-trials variability) compared to the total variance, yields the intraclass coefficient. Three variances are required for computing the intraclass correlations, reflecting the hierarchical data structure shown in Figure 48: (a) variances for measures taken within the same trial at each Peak-to-Peak Segment, (b) variances for measures across trials produced by each individual participant, and (c) total variance across all the data in the study.

Table 6 shows the standard deviations for subjects, trials, and for residuals from measurements taken at each peak. The values of the intraclass correlations, $ICC_{participant}$ and ICC_{trial} are calculated from these standard deviations, with formulas obtained from West et al. (2007) and Zuur, Ieno, Walker, Saveliev, and Smith (2009) shown in Formulas 20 and 21 on the following pages.

Table 6

Standard Deviations for Participant, Trial, and Residuals for Spatial Errors (i.e., log (RMSE) Used in ICC Computations

Coefficient	Standard deviation (σ)
Participants (estimate of σ_a)	0.1637
Trials nested in participants (estimate of σ_b)	0.3426
Residuals of measurements nested in trials (Measures at peaks within samples, estimate of σ_ε)	0.4244

As shown in Table 6, the estimated value of σ_ε is 0. 4244; for σ_a, it is 0.1637; and for σ_b, it is 0.3426. To obtain the intraclass correlation for participants, these values are substituted in Formula 16, presented previously and reproduced here, and Equation 32, respectively.

For participants,

$$ICC_{participant} = \frac{\sigma^2_{participant}}{\sigma^2{}_{participant} + \sigma^2_{trial} + \sigma^2} \qquad \text{(Formula 16)}$$

The random variation for participants is depicted in Table 6 and Formula 16 by σ_a^2, represented by $\sigma^2_{participant}$ and σ_b^2, represented by σ^2_{trial}. The total random variation is represented by the denominator in Formula 16. Using the standard deviations from Table 6, the calculation for the intraclass correlation for participants is shown by Equation 32.

Intraclass correlation between participants was calculated using Equation 32, substituting $\sigma_{participant}$ for σ_a, and σ_{trial} for σ_b:

$$ICC_{participant} = \frac{\sigma^2_{participant}}{\sigma^2_{participant} + \sigma^2_{trial} + \sigma^2} = \frac{0.1637^2}{0.1637^2 + 0.3426^2 + 0.4244^2} = 0.0804$$

(Equation 32)

$ICC_{participant}$ was relatively low (see also the box-and-whisker plots in Figures 49-50), suggesting that individual participants produced a wide range of responses. These results provide additional support for the decision to model participants as random effects. A closer look at individual participants was warranted to determine whether the variability was random or systematic, e.g., low within experimental conditions, but high across conditions.

As described in previous sections, intrasubject variability within trials can be assessed by the formula for ICC_{trial} (Formula 17), which produces a numerical value for the correlation of the 11 observations of the same trial.

As presented previously, for trials,

$$ICC_{trial} = \frac{\sigma^2_{participant} + \sigma^2_{trial}}{\sigma^2_{participant} + \sigma^2_{trial} + \sigma^2}$$ (Formula 17)

Intraclass correlation between individual spatial errors within the same trial by the same participant was calculated using Equation 33.

$$ICC_{trial} = \frac{\sigma^2_{participant} + \sigma^2_{trial}}{\sigma^2_{participant} + \sigma^2_{trial} + \sigma^2} = \frac{0.1637^2 + 0.3426^2}{0.1637^2 + 0.3426^2 + 0.4244^2} = 0.4446$$

(Equation 33)

ICC_{trial} was considerably higher than $ICC_{participants}$, indicating that participants were moderately consistent in their responses *within* each trial, even though their responses varied *between* trials, particularly when performed under different experimental conditions.

4.9.8. Final comments on the model for analyzing spatial errors

In summary, the final model for spatial errors, RMSE, reproduced in Equation 31 appears to fit the data adequately.

$$Log\,(RMSE_{ijk}) = Feedback_{ijk} + TargetRate_{ijk} + Feedback_{ijk} * TargetRate_{ijk} + f_{ft}(Peak) + a_i + b_{j|i} + \varepsilon_{ijk}$$

(Equation 31)

The residual plots in Figures 62 and 63, and the histogram of the residuals in Figure 65 suggest that the GAMM with random effects consisting of participant-specific and trial-specific intercepts, nested in participants, with varIdent structure for Feedback was effective in reducing the heteroscedasticity in the first linear mixed effects model for the RMSE values. Thus, a generalized additive mixed modeling procedure appeared to be an adequate choice for analyzing spatial errors, and effectively accounted for the subject-specific effects, correlated observations within trials and participants, heteroscedasticity, and nonlinearity that characterized the spatial error data.

In the next chapter, I will outline the specific results of each statistical analysis to determine the effect of Feedback type, Target Rate, Gender, and Peak-to-Peak Segment on the wiggle, timing, and RMSE values obtained in this study.

CHAPTER V: RESULTS

5.1 Overview

This chapter addresses the questions that motivated the present research, beginning with a discussion of the methods used to test spatial accuracy. Questions pertaining to timing accuracy and smoothness in the writing line are then addressed. The model used to analyze the responses and provide statistical analyses for each dependent variable is described. Finally, results of statistical analyses are presented to determine whether the results support or fail to support the hypotheses presented in Chapter 1.

5.2. Results of the Spatial Error Analysis

5.2.1. Spatial Errors

The spatial accuracy methods were developed primarily to validate the ability of the software to quantify the dependent measures in this study. Reliable software should produce results that are consistent with robust findings on spatial features of movement. Two well-documented spatial characteristics of movement were selected as a test of the software:

(1) Participants should trace the shape of the target more accurately when they can see how closely their pen movements match the path of the target (Guadagnoli & Kohl, 2001). Thus, participants are expected to copy the target path more accurately when the tracks of their pen movements are visible than in trials where the pen movements do not leave a visual track.

(2) Spatial accuracy and timing accuracy between the computer target and pen movements should be negatively correlated, in accordance with Fitts' Law, which predicts a tradeoff between movement speed and spatial accuracy (Fitts, 1954; Bosga, Meulenbroek & Rosenbaum, 2005; Zhai, Kong, & Ren, 2004). Since support for these hypotheses was expected, the results could be used to validate the subroutines used in the present study to measure participant responses in the present study.

Plots of spatial errors for each peak-to-peak segment, determined as the root mean square error (RMSE) of the differences between the curves drawn by the target and those drawn by the participants, revealed a sizable gap between Feedback:Absent and Feedback:Present conditions (Figure 67).

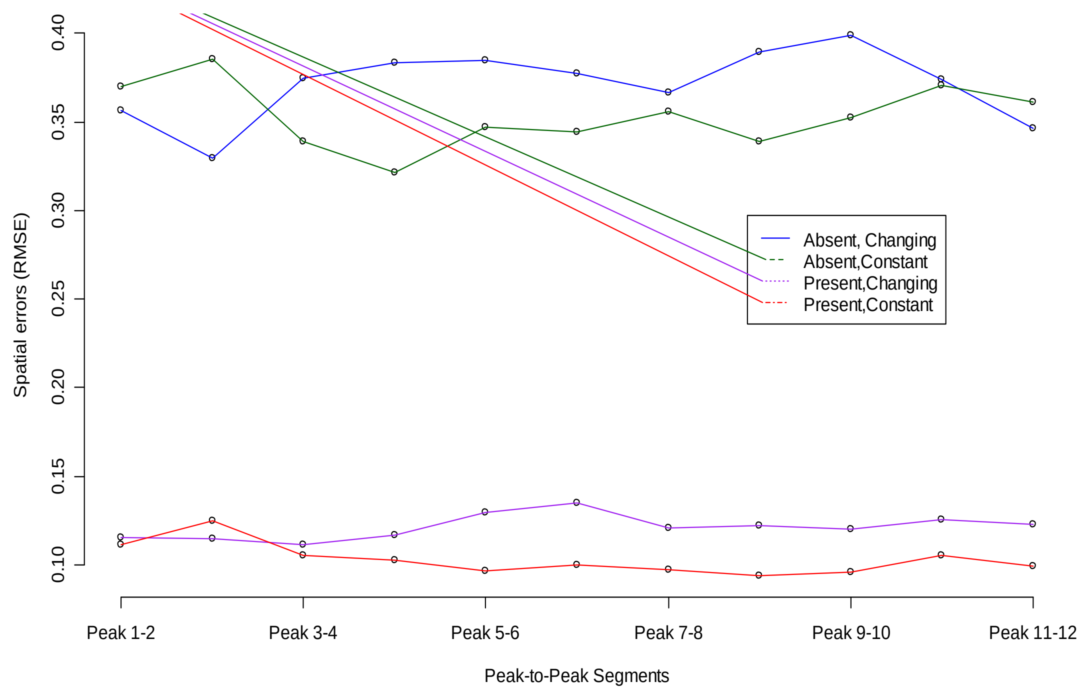

Figure 67. Root Mean Square Error (RMSE) values (*n* = 532-540 scores at each point) for spatial errors at each peak-to-peak segment under four experimental conditions, i.e.,

(a) Feedback:Absent-TargetRate:Changing, (b) Feedback:Absent-TargetRate:Constant,

(c) Feedback:Present-TargetRate:Changing, and (d) Feedback:Present-TargetRate:Constant conditions.

5.2.2. Optimal Model for Spatial Errors

Spatial errors (RMSE) were analyzed using a generalized additive mixed model (GAMM) with smoothers for each Feedback:TargetRate combination. *F*-tests were used to test the null hypothesis that the estimated percentages of RMSE of the fixed factor ($\hat{\beta}$) equaled a hypothesized value (β_0), i.e., H_0: $\hat{\beta} = \beta_0$. The log of the response [i.e., log(RMSE)] was used in the final optimal model, which contained the fixed factors of Feedback, TargetRate, and the interaction between Feedback and Target Rate (denoted by Feedback$*$TargetRate), with a smoothed Peak term that allowed peak-to-peak segments to have different values, depending on the levels for Feedback (Absent or Present) and Target Rate (Changing and Constant). As a result, the optimal GAMM included four smoothers, one for each Feedback and Target Rate condition, as shown by Equation 31, reproduced here.

$$log(RMSE_{ijk}) = Feedback_{ijk} + TargetRate_{ijk} + Feedback_{ijk} * TargetRate_{ijk} + f_{ft}(Peak) + a_i + b_{j|i} + \varepsilon_{ijk}$$

with a varIdent covariance structure for Feedback (Equation 31)

Participants were entered into the model as random effects, and trial level (1-16 for each experimental session) was nested in participants, allowing random intercepts for each participant and each trial within the corresponding participants. A *varIdent* structure was imposed on Feedback, allowing different variances for Feedback:Absent and Feedback:Present conditions.

Trial order did not have an effect on spatial errors: neither trial order over the experimental session (F = 0.656, p = 0.418) nor trial order within each separate experimental condition (F = 0.1237, p = 0.9461) was significant. Gender was not found to be a significant predictor of spatial errors (F = 0.546, p = 0.46), and no significant interaction between gender and other predictors was found. Therefore, trial order and gender were not included in the optimal model.

5.2.3. Statistical Analysis of Spatial Errors

5.2.3.1. Parameter Estimates and Statistical Tests

Results of statistical analyses indicated that spatial errors were smaller when the traces of the participants' pen movements were presented on the screen in the same space as the moving target. Table 7 includes parametrically estimated values for Feedback, Target Rate, and the interaction between Feedback and Target Rate (i.e., Feedback∗TargetRate) that "enter the model as non-parametrically smoothed functions" (Crawley, 2007, p. 611) and are estimated by the generalized cross validation criterion which determines the optimal smoothing function (Zuur et al, 2009). Parameters for the effects of Feedback, Target Rate, and Feedback∗TargetRate on RMSE (Table 7) showed statistically greater RMSE when Feedback was absent. The effects of Feedback and Target Rate were interdependent, i.e., the effect of Feedback on spatial error was dependent on whether Target Rate was changing or constant, and the effect of Target Rate was dependent on whether Feedback was absent or present. Note that standard errors (SE) in Table 7 pertain to contrasts, not to the estimates themselves (Crawley, 2007). *t*-tests were used to test the null hypothesis that the estimated RMSE of the fixed factor ($\hat{\beta}$) equals a hypothesized value (β_0), i.e., H_0: $\hat{\beta} = \beta_0$.

Table 7

Estimated Values and Statistical Significance for Predictors (Feedback and TargetRate) in Optimal Model for log(RMSE)

Predictor	Estimated RMSE	SE	*t*-value	*p*-value
Intercept	-1.3517	0.0394	-34.291	<0.0001
Feedback:Present	-0.6917	0.0348	-19.895	<0.0001
TargetRate:Constant	0.1467	0.0428	3.429	0.0006
Feedback:Present-TargetRate:Constant	-0.4910	0.0568	-8.642	<0.0001

The summary (model) command in the statistical program *R* produced a table that used the first factor (alphabetically) in the model for the *Intercept*. Estimates for subsequent factors were the differences between the intercepts for those factors and the intercepts for the first group (Zuur et al., 2009, p. 336). Thus, the RMSE in the Feedback:Absent-TargetRate:Changing condition was -1.3517. Estimated values for each term (Feedback:Present-TargetRate:Constant, and Feedback:Present-

TargetRate:Constant) were added to the intercept value for the first group (Feedback:Absent-TargetRate:Changing). The estimated value for RMSE in the Feedback:Present condition was -0.6917; therefore, the log(RMSE) in the Feedback:Present-TargetRate:Changing condition was lower than in the Feedback:Absent-TargetRate:Changing condition (i.e., -1.3517 -0.6917 = -2.0434). The estimated RMSEs in the Feedback:Absent-TargetRate:Constant condition were 0.1467 higher than in the Feedback:Absent-TargetRate:Changing condition (i.e., -1.3517 + 0.1467 = -1.205), but when Feedback was present in the constant rate condition (Feedback:Present-TargetRate:Constant), the RMSE was lowest (-1.3517 + -0.6917 + 0.1467 + -0.4910 = -2.3877). To obtain the fitted value at any particular peak, the contribution of the smoother at the peak of interest was added to these initial values for each condition. An *F*-test indicated that each smoother was significant ($p < 0.0001$). These results supported the observation that the relationship between the Feedback, Target Rate and spatial errors was nonlinear. As shown in Table 7, the fewest spatial errors, as indicated by the lowest log(RMSE) scores, occurred in the Feedback:Present-TargetRate:Constant condition. The greatest spatial errors, indicated by the highest RMSE scores, were observed in the Feedback:Absent-TargetRate:Constant condition.

Table 8

Estimated Mean Effects of Predictors (Feedback and TargetRate) in Optimal Spatial Error log(RMSE) Model

TargetRate	Feedback:Absent	Feedback:Present	Mean
changing	-1.1776 ± 0.0162	-2.2275 ± 0.0118	-1.7287 ± .01385
n[a]	(1515)	(1516)	(3031)
constant	-1.205 ± 0.0158	-2.3881 ± 0.0119	-1.8133 ± .01473
n	(1409)	(1480)	(2889)
Mean	-1.1930 ± .0113	-2.307 ± -.0085	-1.7501 ± 0.0102
n	(2924)	(2996)	(5920)

[a] n is the number of peak-to-peak segments measured for each condition

The fitted value at each peak was obtained by adding the contribution of the relevant smoother (Feedback-TargetRate) for each peak (Chapter 4, Table 5).

5.2.3.2. Effects of Predictors in Interactions

The optimal model contained the significant interactive term, Feedback∗TargetRate, precluding interpretation of Feedback and Target Rate as separate main effects on the RMSE data since "[i]interpretations of main effects are impossible, or at least, unreliable when there is an interaction" (Underwood, 1997, p. 321). However, the effects of individual factors could be estimated using likelihood ratio tests to compare models that differ by only the factor of interest, i.e., by within nested ANOVA deletion tests (Crawley, 2007). In the present study, a model containing the factor of interest and all other significant factors was compared to a second model lacking only the factor of interest, as both a separate variable and as an element of an interaction. The results showed that the final model containing the Feedback∗TargetRate interaction provided a better fit to the data than models without Feedback or without Target Rate (Table 9). Thus, Feedback and Target Rate, considered individually, are highly significant predictors of spatial error.

Table 9

Comparison of Models Deleting Predictors in Feedback∗TargetRate interaction from Optimal Spatial Error Log (RMSE) Model

	Model	df	AIC	BIC	logLik	Test	L.Ratio	*p* value
Feedback:Target Rate	1	14	7458.632	7552.237	-3715.316			
Feedback only	2	11	7505.495	7579.042	-3741.747	1 vs. 2	52.8628	<.0001
Target Rate only	3	11	7922.866	7996.413	-3950.433	1 vs. 3	470.2335	<.0001

In Table 9, *AIC* is a measure of the complexity of the model: the lower the value, the better the model fits the data. *logLik* represents the value of the logarithm of the likelihood, or the product of the values of the probability density function for a given set of data, used in Likelihood Ratio (*L.Ratio*) tests. Test refers to the models compared, e.g., comparison of Model 1 and Model 2. The Likelihood Ratio value was compared to a distribution to determine the likelihood that the models being compared provide statistically similar fits to the data. The *p*-value indicates the probability that the models compared would produce such extreme Likelihood Ratios, if the null hypothesis were true, i.e., if there was no significant difference between models.

The results in Table 9 show that the final model containing the Feedback∗TargetRate interaction was a better fit to the data than models without Feedback (L = 2298.033, p <.0001) or without Target Rate (L = 71.504, p <.0001). Thus, even though we cannot interpret the variables Feedback and Target Rate directly as main effects due to the interaction in the final model, the process of comparing nested models with a deletion test indicated that Feedback and Target Rate, evaluated singly, were highly statistically significant predictors of spatial errors in this study.

5.2.3.3. Intraclass Correlations

Similarities in the responses among subgroups (i.e., within participants and within trials) were determined using intraclass correlations. Intraclass correlations for participants, $ICC_{participants}$, (0.0804) was relatively low, corresponding to the box-and-whisker plots in Figures 52, 53, and 54, in Chapter 4, *Statistical Analysis*, suggesting that each participant demonstrated a wide range of responses. A closer look at individual responses was obtained by looking at ICC_{trial}, i.e., the intraclass correlation between peak-to-peak errors within the same trial by the same participant. In the spatial error data, ICC_{trial} (= 0.4446) was considerably higher than $ICC_{participants}$, indicating that participants were moderately consistent in their responses within each trial, meaning that they produced similar peak-to-peak copying errors across the peaks in a sample, but values of the RMSEs varied from trial to trial.

5.3 Tests of Spatial Accuracy Hypotheses

5.3.1 Spatial Accuracy with Concurrent Visual Feedback

The first spatial accuracy hypothesis proposed that spatial accuracy with concurrent visual feedback would be superior to spatial accuracy without visual feedback. Results of the tracking trials supported the hypothesis that concurrent visual Feedback fostered spatial accuracy between the participant loop-drawing response and the animated target. Better spatial matching occurred when the visual traces of the participant drawings were displayed on the screen along with the target, as indicated by lower RMSE values in both Feedback:Present conditions (i.e., Feedback:Present – TargetRate:Changing and Feedback:Present – TargetRate:Constant) (Tables 10 and 11). This expected result verified that the software was performing as intended.

5.3.2. Drawing Speed

The second spatial hypothesis predicted that spatial errors would increase with faster movement speeds. As with the first spatial hypothesis, the goal of testing this hypothesis was to verify that the software was performing as intended in tracking timing and spatial errors, since speed accuracy-tradeoffs are commonly observed. However, the correlation between sample speeds and spatial errors could not be tested under all conditions, since mean sample speed operated as both a predictor and an outcome of spatial accuracy conflating predictors and responses (Gary Oehlert, University of Minnesota School of Statistics, personal communication, June 11, 2008). Such correlations between predictors and errors are termed *endogeneity,* and produce biased and inconsistent estimates and misleading results (Shepherd, 2008). Endogeneity could be expected under three experimental conditions in the present study: (a) Feedback:Absent-TargetRate:Changing, (b) Feedback:Present-TargetRate:Changing, and (c) Feedback:Present-TargetRate:Constant. Endogeneity is less likely to influence results in the Feedback:Absent-TargetRate:Constant condition where the participants' visual traces were not presented and the Target Rate remained consistent over the sample. Therefore, the second spatial hypothesis, which tested the relationship between spatial errors and increased movement speed, was tested only under the Feedback:Absent-TargetRate:Constant condition.

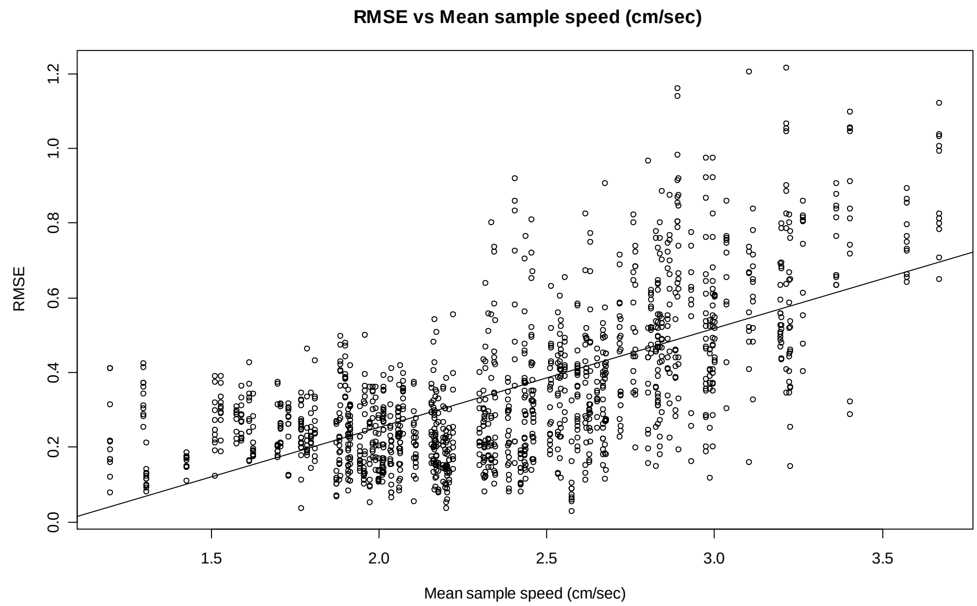

Figure 68. Spatial error (RMSE) at different mean sample speeds (cm/sec) produced by participants in the Feedback:Absent-TargetRate:Constant condition. Each point represents the RMSE value for one peak-to-peak segment, plotted against the mean sample speed for the corresponding segment.

The correlation between pen movement speed and spatial errors (RMSE) in the Feedback:Absent-TargetRate:Constant condition shown in Figure 68 was positive and significant ($r = 0.648$, $p < 0.0001$), indicating that faster speeds were associated with increased spatial errors, indicated by higher root mean square error (RMSE) values for the participant drawings in relation to the target curves of the stimulus. Thus, the results in the Feedback:Absent-TargetRate:Constant condition appeared to support the second spatial hypothesis regarding a speed-accuracy trade-off, although this result must be interpreted cautiously due to the possible effects of endogeneity. These results, however, are consistent with a large body of literature, and suggest that the software and the participants performed as intended.

5.4. Results of the Timing Error Analysis

5.4.1. Timing Errors

To begin the data analysis, plots of timing errors at each peak along the sample trajectory were created. Timing errors were defined as the difference (in seconds) between the times that the animated target and the participant reached corresponding loop peaks along the sample (see Chapter 4 for details on the difference measures). The plots revealed that fewer timing errors occurred in Feedback:Absent conditions than in Feedback:Present conditions, with the fewest timing errors found in the Feedback:Absent-TargetRate:Changing condition (Figure 69).

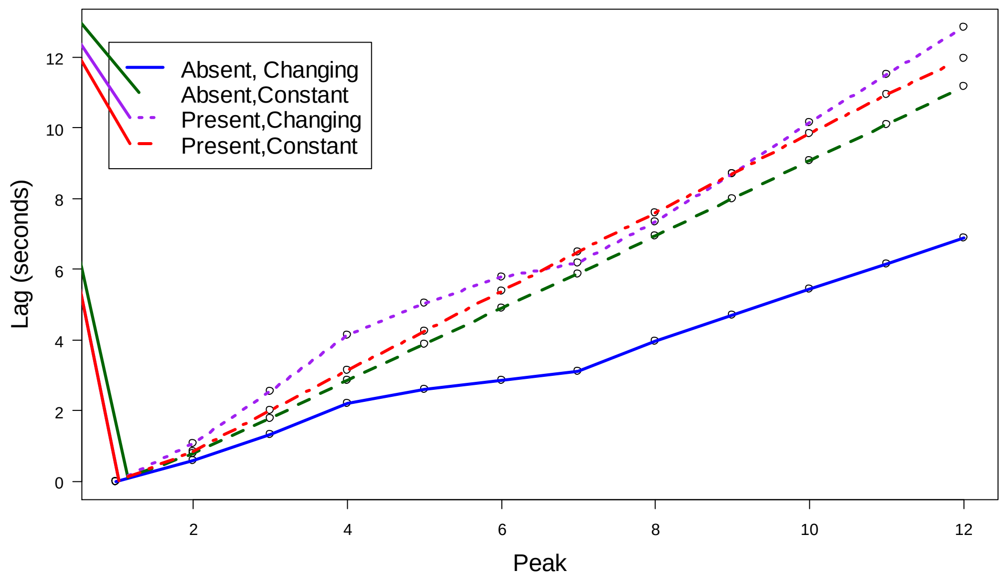

Figure 69. Timing differences (seconds, *n* = 130-140 scores at each point) for timing errors at each peak (1-12) in the sample under four experimental conditions:

(a) Feedback:Absent-TargetRate:Changing, (b) Feedback:Absent-TargetRate:Constant,

(c) Feedback:Present-TargetRate:Changing, and (d) Feedback:Present-TargetRate:Constant.

5.4.2. Optimal Model for Timing Errors

Timing errors were analyzed using a linear mixed model with random slopes for each peak in the sample, participant specific intercepts (i.e., random intercepts) and a varIdent covariance structure for treatments, which allows each type of treatment (i.e., the conditions formed by differing Feedback*TargetRate combinations) to have a separate variance.

The final optimal model contained fixed effects of the single predictors of Feedback, TargetRate, Gender, and Peak; the two-way interactions between Feedback and TargetRate (Feedback*TargetRate), Feedback*Gender, Feedback*Peak, and Gender*Peak; and the three-way interactions among Feedback*TargetRate*Peak and Feedback*Gender*Peak, as shown by Equation 22, reproduced here.

$$\begin{aligned}
&\mathit{TimingError}_{ijk} = \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk}*\mathit{Peak(centered)}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Gender}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Gender}_{ijk} + \mathit{Feedback}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{TargetRate}_{ijk}*\mathit{Peak(centered)}_{ijk} + \mathit{Gender}_{ijk}*\mathit{Peak(centered)}_{ijk} \\
&+ \mathit{Feedback}_{ijk} + \mathit{Target}_{ijk} + \mathit{Gender}_{ijk} + \mathit{Peak(centered)}_{ijk} + \varepsilon_{ijk} \\
&\textit{with varIdent covariance structure for } \mathit{Feedback}_{ijk}*\mathit{TargetRate}_{ijk}
\end{aligned}$$

(Equation 22)

Participants were entered into the model as random effects, and trial level (1-16 for each experimental session) was nested in participants, allowing random intercepts for each participant, with a varIdent covariance structure. Random slopes were allowed for peaks.

Trial order was not found to have an effect on timing errors: neither trial order over the experimental session (F = 0.033, p = 0.855) nor trial order within each separate experimental condition (F = 0.81, p = 0.4864) was significant. Therefore, trial order was not included in the optimal model.

5.4.3. Statistical Analysis of Timing Errors

5.4.3.1. Parameter Estimates and Statistical Tests

The statistical analyses revealed that timing errors were greater when the traces of the participants' pen movements were presented on the screen concurrently and within the same space as the moving target stimulus where the mean timing errors in the Feedback:Present conditions (upper panels of Figure 70) are higher than in the Feedback:Absent conditions (lower panels).

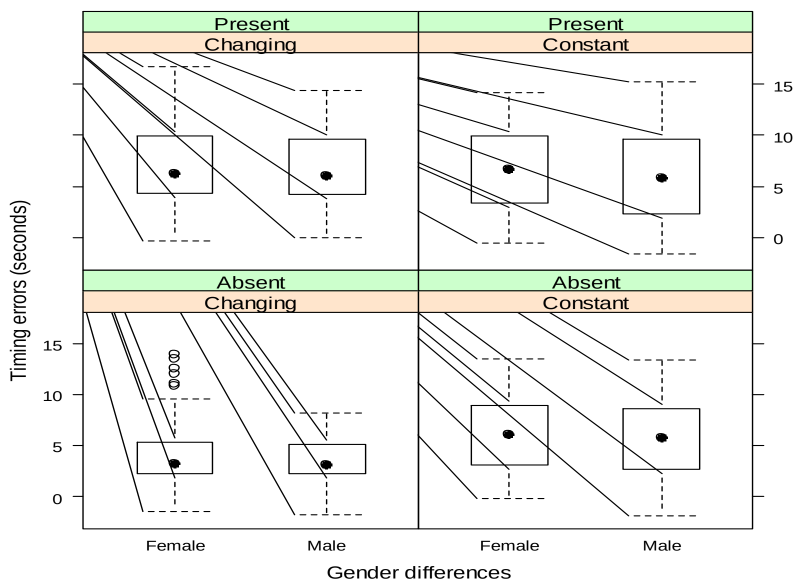

Figure 70. Mean timing differences in each experimental condition (*n* = 1439-1540; designated by gender (females: *n* = 957-1010; males: *n* = 470-528). The black dots in the boxplots indicate the median value in each category. The box is bounded by horizontal lines at the 25^{th} and 75^{th} quartile, allowing the box to represent the spread of the central 50% of the data. "Whiskers" extend 1.5 times the interquartile range (the

length of the box, between the 25th and 75th percentile). The dots represent outliers, i.e., values that are greater than 1.5 times the interquartile range.

Table 10 includes parametrically estimated values for Feedback, Target Rate, Gender, Peak and the significant interactions among these variables. These parameters showed statistically greater timing errors when Feedback was Present. The effects of Feedback and Target Rate were interdependent, i.e., the effect of Feedback on timing errors was dependent in part on whether the Target Rate was changing or constant, and the effect of Target Rate was dependent on whether Feedback was absent or present. Timing errors increased with each successive peak. Interactions showed that timing errors by males were reduced when feedback was absent, but increased when they observed their own performance in the feedback present condition.

Table 10

Estimated Values and Statistical Significance for Predictors (Feedback, Target Rate, Peak. Gender, and their interactions) in Optimal Model for Timing Errors

Predictor	Est.Timing Error	SE	df	*t*-value	*p*-value
(Intercept)	2.5024	0.0838	5414	29.872	<0.0001
Feedback(FB)Present	2.2062	0.0972	504	22.7035	<0.0001
TargetRate(TR)Constant	1.4320	0.0859	504	16.6793	<0.0001
Peak-centered(c)	0.6089	0.0195	5414	31.3029	<0.0001
GenderMale	-0.2347	0.1241	33	-1.8917	0.0673
FBPresent:TRConstant	-1.8304	0.1220	504	-14.9994	<0.0001
FBPresent:Peak(c)	0.5353	0.0219	5414	24.4308	<0.0001
TRConstant:Peak(c)	0.4380	0.0192	5414	22.8084	<0.0001
FBPresent:GenderMale	0.0140	0.1284	504	0.1087	0.9135
Peak(c):GenderMale	-0.0292	0.0291	5414	-1.0051	0.3149
FBPresent:TRConstant:Peak(c)	-0.4289	0.0274	5414	-15.6570	<0.0001
FBPresent:Peak(c):GenderMale	-0.0807	0.0288	5414	-2.8035	0.0051

Note that standard errors (SE) in Table 10 pertain to contrasts, not to the estimates themselves (Crawley, 2007). *t*-tests were used to test the null hypothesis that the estimated timing errors of the fixed factor ($\hat{\beta}$) equals a hypothesized value (β_0), i.e., H_0: $\hat{\beta} = \beta_0$.

The summary (model) command in the statistical program *R* produced a table that used the first factor (alphabetically) in the model for the *Intercept*. Estimates for subsequent factors were the differences between the estimates for the *Intercept* and the intercepts of other fixed factors in the model (Zuur, 2009). Thus, the intercept for the timing error data in the Feedback:Absent-TargetRate:Changing condition for females (2.5024) was listed first, because it was alphabetically the first combination for the Feedback∗TargetRate∗Gender*Peak interaction. Estimated values for each term (Feedback:Present, TargetRate:Constant, Peak (centered). Gender:Male, and the remaining interactions were added to the values for the intercept single factors. In each case, I estimated the value where the peak was centered at 0. Peak = 5 was centered at 0 because the statistical program *R* failed to converge at the true center (peak = 6). Thus, the estimated value for timing errors by females at peak 0 under the Feedback:Absent – TargetRate:Changing condition was 2.5024 + .6089 = 3.1113 seconds.

Estimates for other variables at the centered peak were obtained by adding to this basic value. For example, the added timing error (seconds) in the Feedback:Present condition was 2.2062; therefore, the timing errors in the Feedback:Present-TargetRate:Changing-Gender:Female data were on average 5.3175 at the value for which the centered peak = 0. The value is calculated by adding the estimates for

Feedback:Absent-TargetRate:Changing-Gender:Female + Feedback:Present (2.5024 + 6089 + 2.2062 = 5.3175).

Using this system, the greatest timing errors (lags behind the target) occurred in samples drawn by female participants in the Feedback:Present-TargetRate:Constant condition, resulting in a mean lag of 5.8923 seconds at the centered peak for these samples (p <.0001). In general, timing errors were greater in the Feedback:Present conditions and TargetRate:Constant conditions, although the greatest effect was the presence of Feedback. Further, Males were likely to have lower timing errors than Females in each Feedback-TargetRate condition. The means for each Feedback*TargetRate*Gender grouping at the centered Peak measure are listed in Table 11, below.

Table 11

Estimated Mean Effects of Predictors (Feedback∗TargetRate*Gender) in Optimal Timing Error Model

Feedback	TargetRate: Changing	TargetRate: Constant	Gender Female	Gender Male	Mean ± SE
Absent	3.614 ± 0.053		3.677 ± 0.067	3.492 ± 0.086	
n[a]	(1536)		(1010)	(526)	4.774 ± 0.057
					(2963)
		5.928 ± 0.092	6.074 ± 0.109	5.631 ± 0.167	
n		(1427)	(957)	(470)	
Present	6.843 ± 0.094		6.903 ± 0.118	6.732 ± 0.154	
n	(1518)		(990)	(528)	
					6.713 ± 0.070
					(3001)
		6.460 ± 0.103	6.704 ± 0.120	5.987 ± 0.190	
n		(1483)	(978)	(505)	

[a] *n* is the number of peaks measured for each condition and gender

5.4.3.2. Effects of Predictors in Interactions

The optimal model contained the significant interactive term, Feedback∗TargetRate, precluding interpretation of Feedback and TargetRate as separate main effects for timing errors (Underwood, 1997). The effects of Feedback and Target Rate were estimated using likelihood ratio tests within nested ANOVA deletion tests (Crawley, 2007), as shown in Table 12. The results were compared to the model with the Feedback∗TargetRate and Feedback*TargetRate*Peak interaction, to assess the significance of the Feedback and TargetRate separately. In these data, the optimal model with terms Feedback∗TargetRate*Peak was compared to a model with Feedback only or TargetRate only, to determine whether Feedback and TargetRate contributed significantly to the results.

Table 12

Comparison of Models Deleting Predictors in Feedback∗ TargetRate interaction from Optimal Timing Error Model

	Model	df	AIC	logLik	Test	L.Ratio	*p* value
Feedback∗TargetRate*Gender	1	22	8655.307	-4305.653			
Feedback*Gender only	2	16	9068.708	-4518.354	1 vs 2	425.40	<.0001
TargetRate*Gender only	3	12	10461.570	-5218.780	1 vs 3	1826.26	<.0001

The results of the deletion tests displayed in Table 12 showed that the final model containing the Feedback∗TargetRate*Gender interaction was a better fit to the data than models without Feedback (L = 1826.236 p < 0.0001) or without Target Rate (L = 425.40, p < 0.0001). Thus, even though the variables, Feedback and Target Rate could not be interpreted directly as main effects, comparison of nested models using a deletion test indicated that Feedback and TargetRate, individually, were highly significant predictors of timing errors.

5.5. Tests of Timing Accuracy Hypotheses

5.5.1. Timing Accuracy with Concurrent Visual Feedback

The first timing hypothesis proposed that concurrent visual feedback may improve timing accuracy, as suggested by previously published research on direct perception (reviewed in Chapter 2). The alternate hypothesis proposed that concurrent visual feedback would impede timing accuracy, as suggested by other research results indicating that concurrent visual feedback increases cognitive load.

Results of the tracking trials did not find benefits for concurrent visual feedback. Most observations showed that participants lagged behind the target under all visual conditions, with greater lags occurring in the Feedback:Present conditions, i.e., when the visual trace of the participants' pen movement was presented on the LCD monitor, along with the target. Thus, movement times increased in the presence of concurrent visual feedback, resulting in longer times for completing each sample in Feedback:Present conditions than under Feedback:Absent conditions.

Although participants lagged further behind the target when their visual trace was shown on the screen, the alternative timing hypothesis, i.e., that concurrent visual feedback would function as a distractor and increase task difficulty, could not be tested in the present study, due to methodological limitations. The data could not be used to determine how the visual trace was perceived, so it was not possible to determine whether the presence of the participants' visual trace distracted them from the task or made it more difficult, even though their performance was poorer. In this study, no conclusions can be made to assess the reason for this decrease in performance.

A secondary hypothesis for the effect of concurrent visual feedback on temporal matching, i.e., that improvement in timing accuracy with concurrent visual feedback would be robust when the timing perturbation occurred in the TargetRate:Changing trials, was not supported by the results. On average, timing perturbation produced increased timing errors in the Feedback:Present conditions, but fewer in the Feedback:Absent conditions (Table 10). Thus, the effects of concurrent visual feedback were worsened with the timing perturbation In contrast, the average effect of the timing perturbation enhanced timing accuracy when concurrent visual feedback was not provided.

The third hypothesis regarding timing accuracy predicted that males would have fewer timing errors than females in the tracking tasks, i.e., gender differences would be present, with males demonstrating smaller timing errors due to increased movement speeds. This hypothesis was supported by the results. However, these findings should be considered with caution, due to the small sample size for males.

5.5.2. Consistency in Performance

The hypotheses tested in the present study were designed to assess the effect of concurrent visual feedback on performance, but not to evaluate consistency of performance across trials. Nevertheless, it is interesting to note the apparently high level of timing consistency from trial to trial demonstrated by most participants. Figure 71 shows all sixteen trials under the four experimental conditions for one participant. It is evident that the participant repeated similar timing errors from trial to trial within a narrow range of variability. For most trials, deviations in timing errors at corresponding peaks from trial to trial were low, sometimes as low as several hundredths of a second, even though the trials were randomized. Comparable graphs would show similar consistency for 17 (49%) of the participants in the study.

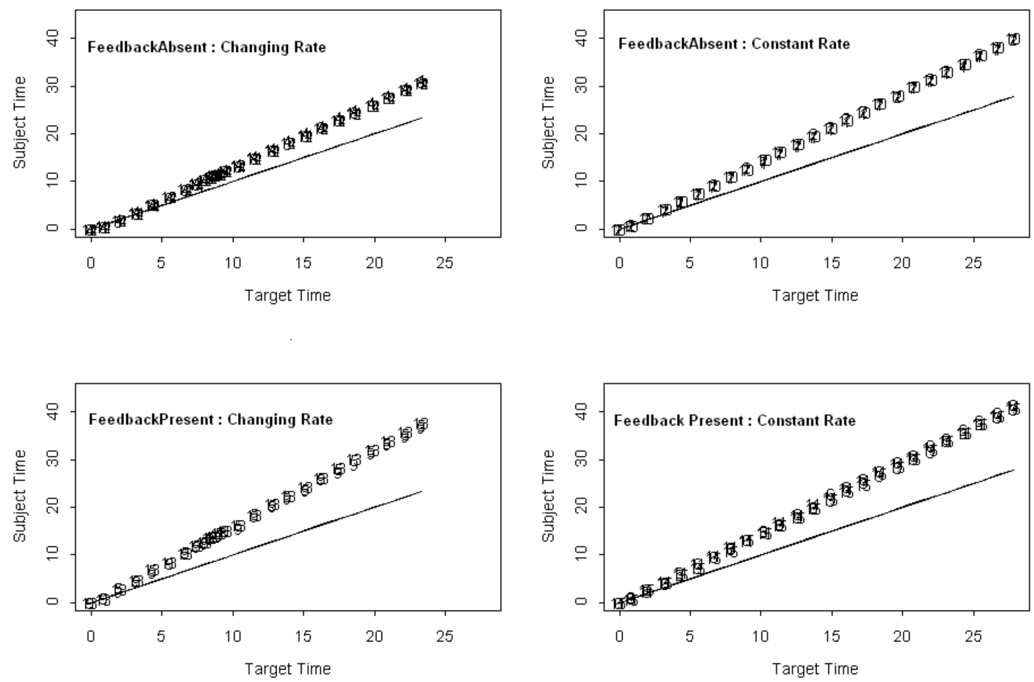

Figure 71. Plots of timing performance in trials under each experimental condition, for one participant who demonstrated consistent performance. The solid line represents the timing of the target. Participant responses on the solid black line would indicate perfect correlation (i.e., frequency coupling) with the target stimulus.

The other eighteen participants (51%) demonstrated trial-to-trial variability in their timing errors during at least one portion of one of the 16 trials. Figure 71 shows results for one of these participants.

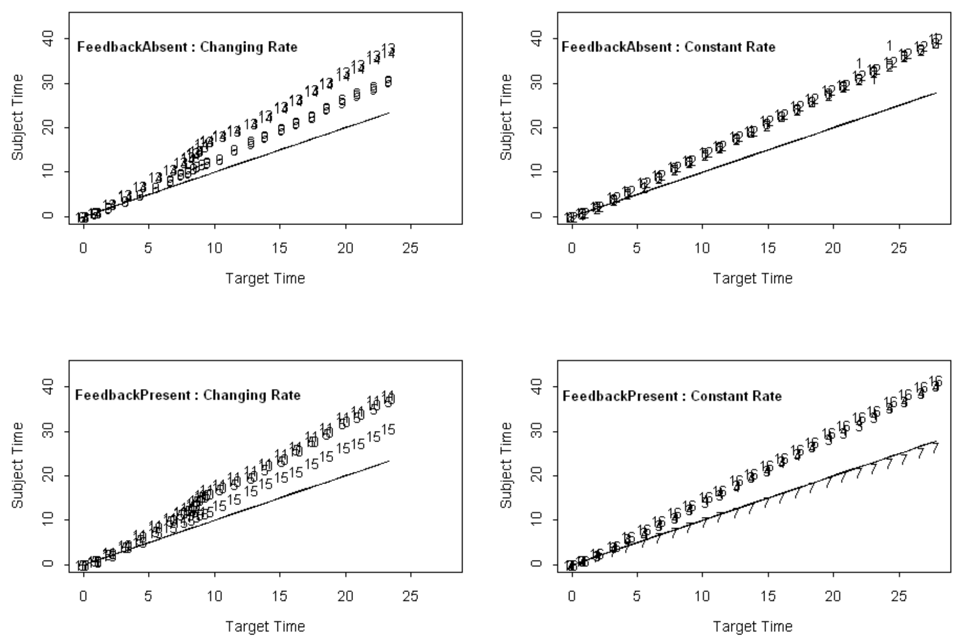

Figure 72. Plots of timing performance by one participant. Each quadrant displays the timing errors for trials performed under each corresponding experimental condition. Although this participant demonstrated inconsistency in each experimental condition for at least two peaks, the majority of the observations are consistent.

Of the 18 participants with inconsistent timing performances, 12 had only one trial that deviated from the timing patterns of all other trials, four participants had two trials that deviated, and two participants had three trials that deviated. Where timing deviations occurred, no trends of increasing or decreasing errors were present, i.e., the participant showed no gradual convergence or increased deviation from the target times in subsequent trials. Thus, the majority of the inconsistent timing errors may have been anomalies, rather than indicators of a changing trend in the response. Further study of the consistency and variability demonstrated by participants is warranted by these observations.

5.6. Results of the Wiggle Analysis

5.6.1. Wiggle

Plots of wiggle, i.e., alternating disturbances in the writing line, were created for each sample under the four experimental conditions. Wiggle was defined as the percentage of the trace that moved in alternating clockwise and counterclockwise directions over a sample, resulting in an undulation that appeared as a zigzag or wrinkle in the writing line (Chapter 4). Plots of wiggle revealed that a lower percentage occurred under Feedback:Absent conditions than under Feedback:Present conditions (Figure 73). The lowest percentage of wiggle was found in the Feedback:Absent-TargetRate:Changing condition. Gender differences were also evident. Males appeared to have less variability under the Feedback:Absent conditions, but greater variability and higher percentages of wiggle under the Feedback:Present conditions.

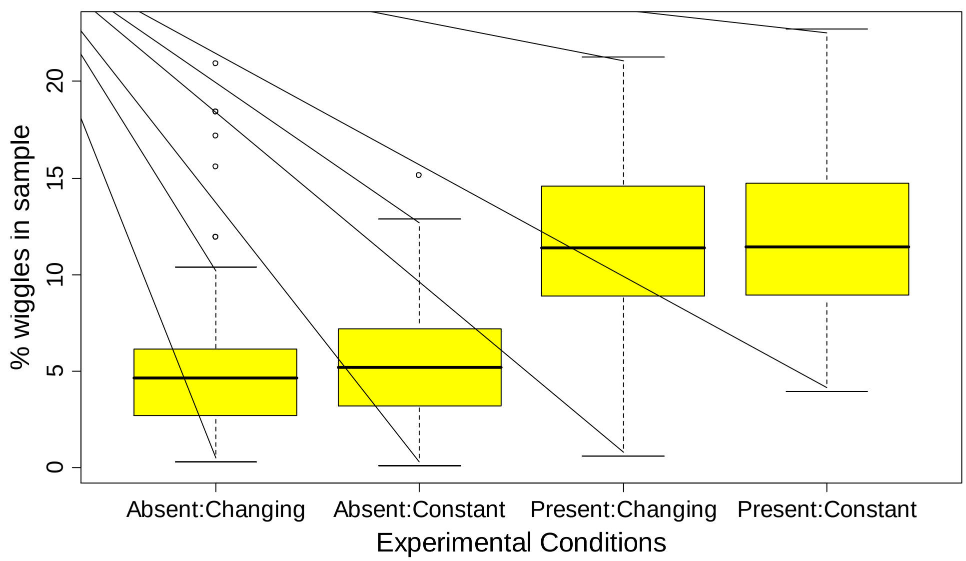

Figure 73. Percentages of wiggle in samples performed under four experimental conditions: Feedback:Absent-TargetRate:Changing, Feedback:Absent-TargetRate:Constant, Feedback:Present-TargetRate:Changing, and Feedback:Present-TargetRate:Constant (*n* = 129-140 scores for each condition). The black dots in the boxplots indicate the median value in each category. The box is bounded by horizontal lines at the 25th and 75th quartile, allowing the box to represent the spread of the central 50% of the data. "Whiskers" extend 1.5 times the interquartile range (the length of the box, between the 25th and 75th percentile). The dots represent outliers, i.e., values that are greater than 1.5 times the interquartile range.

5.6.2. Optimal Model for Percentage of Wiggle in Samples

The percentage of wiggle in each sample was analyzed using a linear mixed model (LME), as described in Chapter 4. *F*-tests were used to test the null hypothesis that the estimated percentages of wiggle of the fixed factor ($\hat{\beta}$) equaled a hypothesized value (β_0), i.e., H_0: $\hat{\beta} = \beta_0$. The final optimal model contained Feedback, Gender, and the interaction between Feedback and Gender, as shown in Equation 15:

$$\textit{Percentage of wiggle}_{ij} = \textit{Feedback}_{ij} + \textit{Gender}_{ij} + \textit{Feedback}_{ij} * \textit{Gender}_{ij} + a_i + b_i + \varepsilon_{ij},$$

with participants as random effects, modeled with participant-specific intercepts and participant-specific slopes for feedback (Equation 15)

Target Rate did not predict percentage of wiggle in the samples (F = 2.0084, p = 0.1571), and there were no significant interactive effects between Feedback and Target Rate (F = 0.1369, p = 0.7115), or Gender and TargetRate (F = 1.7259, p = 0.1896).

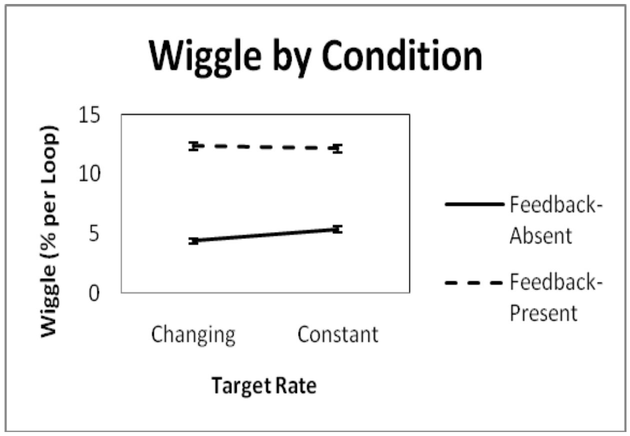

Figure 74. Interaction plot for the effect of Feedback:TargetRate on the percentage of wiggle (n = 129-140 scores for each point).

Trial order did not have an effect on spatial errors: neither trial order over the experimental session (F = 1.4841, p = 0.1066) nor trial order within each separate experimental condition (F =0.2839, p = 0.8370) was significant. Since Trial Order and Target Rate were not significant predictors of wiggle, they were not included in the optimal model.

Participants were entered into the model as random effects, allowing participant-specific intercepts and slopes for feedback, in a random coefficient model with random-intercepts with random-slopes. This model structure resulted in adequately constant variance, allowing the model to use an unstructured covariance structure (i.e., the default covariance structure in *R*).

5.6.3. Statistical Analysis of Wiggle

5.6.3.1. Parameter Estimates and Statistical Tests

The hypothesized factors of Feedback, Gender, and the interaction of Feedback∗Gender jointly affected the percentage of wiggle in the samples. A higher percentage of wiggle was observed under the Feedback:Present conditions than in the Feedback:Absent conditions, but this effect was influenced by the Gender of the participant (Table 13).

Table 13

Estimated Values and Statistical Significance for Predictors (Feedback∗ Gender) in Optimal Model for Wiggle

Predictor	Estimated % of wiggle	SE	df	*F*-value	*p*-value
Intercept	5.0911	0.4118	488	12.3621	<0.0001
Feedback:Present	5.9200	0.4461	488	13.2693	<0.0001
Gender:Male	0.2397	0.7043	33	0.3404	0.7357
Feedback:Present-Gender:Male	2.3250	0.7681	488	3.0271	0.0026

The summary (model) command in the statistical program *R* produced a table that used the first factor (alphabetically) in the model for the *Intercept*. Estimates for subsequent factors were the differences between the estimates for the *Intercept* and the intercepts of other fixed factors in the model (Zuur et al., 2009, p. 336). Thus, the first listed intercept for the wiggle data, Feedback:Absent-Gender:Female (5.0911), indicates that the average percentage of wiggle drawn by females in Feedback:Absent conditions was 5.0911 % of the drawn curve. Estimated values for each additional term were added to the *Intercept* value. For example, the added average percentage of wiggle for Feedback:Present was 5.92%, so the percentage of wiggle for Female participants when Feedback was present was 11.0111% (5.0911% + 5.9200%), which was a statistically significant increase ($p < 0.0001$) over the Feedback:Absent condition.

These results are consistent with the plot showing the effects of the predictors on wiggle, displayed in Figure 75.

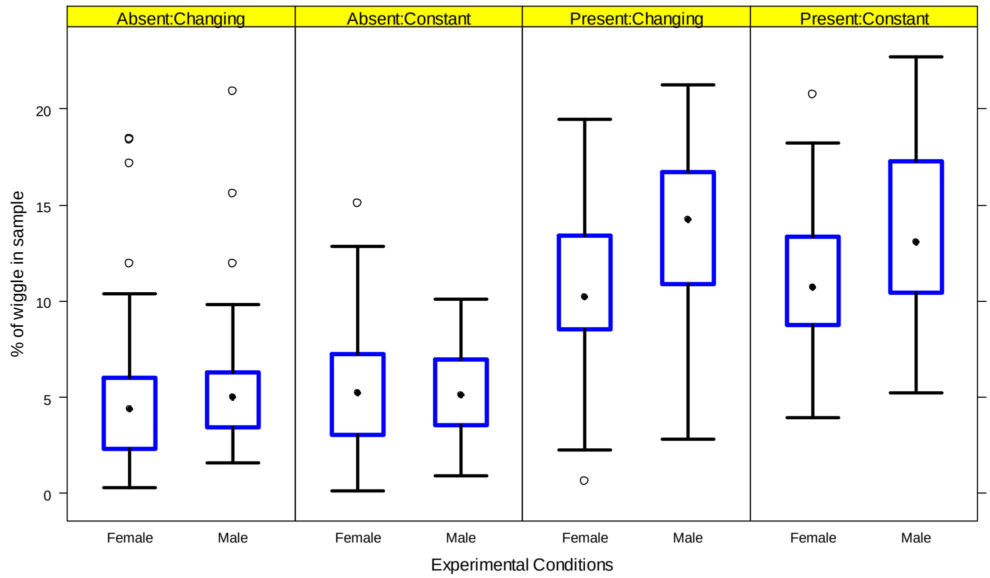

Figure 75. Percentages of wiggle in samples by experimental condition and gender.

The black dots in the boxplots indicate the median value in each category. The box is bounded by horizontal lines at the 25th and 75th quartile, allowing the box to represent the spread of the central 50% of the data. *Whiskers* extend 1.5 times the interquartile range (the length of the box, between the 25th and 75th percentile). The dots represent outliers, i.e., values that are greater than 1.5 times the interquartile range. The mean percentages of wiggle for each gender and experimental condition are presented in Table 14.

Table 14

Estimated Mean Effects of Predictors (Feedback, Gender) in Optimal Wiggle Model

Gender	Feedback:Absent	Feedback:Present	Mean ± SE
Female	5.0911 ± 0.4118	11.0111 ± 0.4461	8.0102 ± 0.4522
n[a]	(178)	(178)	(356)
Male	5.3308 ± 0.7043	13.5758 ± 0.7681	9.4078 ± 0.7759
n	(90)	(86)	(176)
Mean	5.1773 ± 0.4012	11.8668 ± 0.258	8.4966 ± 0.1209
	(268)	(264)	(532)

[a] n is the number of samples measured for each condition.

As shown by Table 14, the percentage of wiggle in the Feedback:Absent condition for male participants (5.3308%) was 0.2397% higher than for females, but the difference was not statistically significant (p = 0.7357). In the Feedback:Present-Gender:Males combination, the average percentage of wiggle (5.0911 + 5.9200 + .2397 + 2.3250 = 13.5758) was higher than for females (11.0111). Thus, effect of concurrent visual feedback during the tracking tasks influenced the percentage of wiggle in samples by males significantly more than for females (p = 0.0026). Overall, the lowest percentages of wiggle were observed in the Feedback:Absent condition with female participants, and the highest percentages were observed in the Feedback:Present condition with males. TargetRates (Changing and Constant) did not produce significant differences in the percentage of wiggle.

5.6.3.2. Effects of Predictors in Interactions

The optimal model for wiggle percentage contained the significant interactive term, Feedback∗Gender, precluding interpretation of Feedback and Gender as separate main effects (Underwood, 1997) on wiggle. The effects of individual factors were estimated using likelihood ratio tests within nested ANOVA deletion tests (Table 15) to compare models that differed by the factor of interest, as both a separate variable and as an element of an interaction (Crawley, 2007). The optimal model with Feedback∗Gender was compared to a model with Feedback only or Gender only, to determine whether Feedback and Gender contributed significantly to the results.

Table 15

Comparison of Models Deleting Predictors in Feedback∗ Gender interaction from Optimal Wiggle Data Model

	Model	df	AIC	BIC	logLik	Test	L.Ratio	*p* value
Feedback:Gender	1	6	2676.219	2710.327	-1330.110			
Feedback only	2	4	2719.517	2736.578	-1355.758	1 vs 2	27.5560	<.0001
Gender only	3	4	3136.992	3154.053	-1564.496	1 vs 3	445.0307	<.0001

In Table 15, *AIC* is a measure of the complexity of the model: the lower the value, the better the model fits the data. *logLik* represents the value of the logarithm of the likelihood, or the product of the values of the probability density function for a given set of data, used in Likelihood Ratio (*L.Ratio*) tests. Test refers to the models compared, e.g., comparison of Model 1 and Model 2. The Likelihood Ratio value is compared to a distribution to determine the likelihood that the models being compared provide statistically similar fits to the data. The *p*-value indicates the probability that the models compared would produce such extreme Likelihood Ratios, if the null hypothesis were true, i.e., if there is no significant difference between models.

As shown in Table 15, the final model containing the Feedback∗Gender interaction provided a significantly better fit to the data than models without Feedback ($L = 445.0307$, $p < 0.0001$) or without Gender ($L = 27.556$, $p < 0.0001$). Thus, even though Feedback and Gender could not be tested directly as main effects, comparison of nested models with a deletion test indicated that, individually, Feedback and Gender are statistically significant predictors of wiggle.

5.6.3.3. Intraclass Correlations

Similarities in wiggle among subgroups (i.e., within trials and within participants) were determined using intraclass correlations, $ICC_{participant}$ and $ICC_{trial.}$ The details of these calculations were presented in Chapter 4, *Statistical Analysis.* Intraclass correlations for participants, $ICC_{participants}$ (= 0.349), was moderate, suggesting variability in each participant's responses over the 16 trials in the experimental session, corresponding to the box-and-whisker plots in Figures 18 and 19 in Chapter 4. A partial explanation for these results may lie in the differences in the participants' responses under the experimental conditions, particularly feedback, where the amount of wiggle was considerably higher in Feedback:Present conditions, compared to trials in Feedback:Absent conditions.

5.7. Tests of Wiggle Hypotheses

5.7.1. Wiggle with Concurrent Visual Feedback

The first wiggle hypothesis proposed that concurrent visual feedback would increase smoothness, reducing the percentage of wiggle in samples drawn with visual feedback when compared to samples drawn without visual feedback. The tempo of the target, moving at a constant or changing rate, was not expected to affect this result. Results of the tracking trials failed to support this hypothesis, as higher percentages of wiggle occurred in the Feedback:Present condition (Tables 15 and 16). The results supported the alternate hypothesis, which proposed that the presence of the visual trace would disrupt the smoothness of the drawing, resulting in increased wiggle. The secondary hypothesis, which predicted that perturbing the target rate would have no significant effect, was supported by the data, since the TargetRate:Constant and TargetRate:Changing conditions had no significant differences.

5.7.2. Gender Effects

The third wiggle hypothesis predicted that gender differences would not be statistically significant, i.e., female and male participants would show similar wiggle responses to feedback. The results did not support this hypothesis. Males exhibited a statistically higher percentage of wiggle than females in all conditions, with greater effects when Feedback was present (Table 15).

CHAPTER VI: DISCUSSION

6.1. Chapter Overview

Chapter 6 begins with a discussion of the results of the study and compares the results to previous research in this area. Next, I describe the limitations of the present study. The chapter will conclude with a description of areas for future research.

The present study examined the effects of concurrent visual feedback in a task of tracking an animated loop-drawing target across an LCD monitor display. Three separate effects were investigated:

(1) Spatial errors (measured by root mean square errors between the size and shape of target loops and participant drawings);

(2) Timing errors (measured according to the times at which the targets and participants reached corresponding peaks of each successive loop in the sample); and

(3) Disturbances in the smoothness of the writing line (measured by the percentage of “wiggles” in the sample).

The robustness of the effect of concurrent visual feedback was assessed by introducing a timing perturbation in eight of the 16 samples provided by each participant.

As described in Chapter 5: *Results*, traces recorded with and without concurrent visual feedback yielded significant differences in spatial errors, timing errors, and disturbances in smoothness in the writing line (wiggles). These results are not surprising since variability in handwriting and tracing tasks performed under different conditions, particularly visual conditions, has been observed frequently, both anecdotally and scientifically. Studies have attributed these differences to various factors, including:

(a) visual and haptic perception (Lagarde & Kelso, 2006); (b) psychological and motor processes (Van Galen & Van Huygevoort, 2000); (c) aesthetics; and (d) semantics, in the case of writing (Doumas & Wing, 2007; Van Galen, Meulenbroek, & Hylkema, 1986).

In the sections to follow, I will discuss the spatial errors, timing errors, and wiggles observed in the present study and compare these findings to previous research in this area.

6.2. Spatial Errors

Two outcomes relevant to spatial errors were observed in the present study:

(1) Spatial accuracy improved with concurrent visual feedback, i.e., presenting the recorded trace of the participants' pen motions in real time in the same space as the animated target stimulus that was being tracked.

(2) An inverse relationship was found between drawing speed and spatial accuracy in the Feedback:Present-TargetRate:Constant condition (the only condition tested for this relationship), consistent with the expectations that more spatial errors would occur at increased speeds.

The effect of concurrent visual feedback in reducing spatial errors is not surprising. The ability of writers to reproduce visual targets more accurately when they can see their graphic output is supported by anecdotal evidence and robust findings throughout the literature (Graham & Weintraub, 1996; Guadagnoli, Leis, Van Gemmert, & Stelmach, 2002; Liu & Scheidt, 2008; Rijntjes, Dettmers, Büchel, Kiebel, Frackowiak, & Weiller, 1999; Roerdink, Ophoff, Peper, & Beek, 2008; Van Galen, Smyth, Meulenbroek, & Hylkema, 1989). Thus, the results of the present study are consistent with previous observations that spatial accuracy in copying tasks improves when the targets and tracings are both visible during the drawing process.

The enhanced spatial accuracy that occurred with visual feedback came at a cost, however. Longer movement times (i.e., slowed movements) occurred with visual feedback, resulting in decreased timing accuracy as participants lagged behind the target. Further, concurrent visual feedback was associated with increased wiggles (decreased smoothness).

6.2.1. Speed-Accuracy Trade-off

A second measure, the inverse relationship between speed and spatial accuracy, was tested only in the Feedback:Absent-TargetRate:Constant condition in the present study due to concerns about the likelihood of endogeneity (conflating predictors with responses) in the other experimental conditions. The goal of this test was to verify that the software was working as intended, since an inverse relationship between accuracy and speed is well documented in literature on movement tasks (Guigon, Baraduc, & Desmurget, 2008; Hancock & Newell, 1985) including those most pertinent to this study: (a) pointing tasks (Fitts, 1954); (b) pursuit tracking (MacKenzie & Ware, 1993; Viviani, Campadelli, & Mounoud, 1987); (c) handwriting (Longstaff & Heath, 1997); (d) curve drawing (Viviani & Terzuolo, 1982); and (e) ellipse drawing (Bosga, Meulenbroek & Rosenbaum, 2005). The relationship between speed and accuracy in the present study was consistent with the aforementioned studies, finding that faster drawing speeds in the Feedback:Absent-TargetRate:Constant condition were associated with increased spatial errors. Thus, the results of spatial errors in the present study provided validation that the software was functioning as intended.

6.3. Disturbances in the Writing Line, or "Wiggle"

In the present study, wiggles were defined as straight, stiff, alternating undulations in the recorded trace. Unlike spatial and timing measures, participants were not informed of wiggle in the writing line as an outcome measure. Wiggles are therefore presumed to be an involuntary response to the experimental tasks. Increased percentages of wiggles were observed in the Feedback:Present conditions when the visual trace of the pen movements was presented concurrently with the moving target. Samples performed in the Feedback:Present conditions were also performed more slowly than those in the Feedback:Absent conditions, suggesting that the visual feedback hindered movement speed. Thus, it is plausible to assume that increased percentages of wiggle may be associated with slower movement speeds. Previous research on disturbances in the writing line has found that decreased smoothness (i.e., increased wiggle) is associated with slower movement speeds and increased biomechanical stiffness (Meulenbroek, Van Galen, Hulstijn, Hulstijn, & Bloemsaat, 2005). For example, Chen, Cha, Chee, and Tappert (2003) found an increase in wrinkliness (analogous to wiggle) at slower movements speeds. In a study of pointing by neurologically intact participants, Keogh, Morrison, and Barrett (2004) found that visual feedback increased biomechanical stiffness and finger tremors, resulting in decreased smoothness in the pointing tasks, and attributed these changes to visual feedback during performance. Their observations suggest that concurrent visual feedback places increased demands on the perceptual-motor system which may result in a higher frequency of movement disturbances. Several studies provide indirect support for this hypothesis, finding that increased task difficulty in writing tasks increases cognitive load and produces increased biomechanical stiffness and axial pressure (Van Den Heuvel, van Galen, Teulings, & Van Gemmert, 1998; Van Gemmert, 1997; Van Gemmert & Van Galen, 1997; and Wing & Baddeley, 1979) and other physiological responses (Fairclough, Venables, Tattersall, 2005). These findings suggest that concurrent visual feedback of the recorded trace may have placed higher demands placed on the perceptual-motor system and contributed to a higher frequency of wiggles by increasing cognitive load and biomechanical stiffness while reducing movement speeds.

6.4. Timing Errors

6.4.1. Effects of Visual Feedback on Timing Errors

Concurrent visual representation of the participants' pen motions with the target (i.e., the Feedback:Present conditions) improved spatial accuracy but impeded timing accuracy by increasing the delay between the target action and corresponding hand motions. In the majority of the timing measures (98.6%), participants lagged behind the target, increasing their lag as they traced across successive peaks across the sample. In several cases where the participant movements were faster than the rate of the target, participants commented later that they questioned the instructions, believing instead that the true intent of the experiment was to see whether participants could race the target to the goal. Thus, it is not possible to interpret these responses as examples of superior temporal coupling.

The extent and consistency of timing lags observed in the present study conflict with findings of several other studies on the effect of vision in coordinated activity, where vision was found to enhance timing accuracy, in general (Bardy & Laurent, 1998; Keogh, Morrison, & Barrett, 2004; Schmidt & Turvey, 1994; Turvey, Shaw, & Mace, 1978). Of particular relevance to the present study, several studies have shown the benefits of concurrent visual feedback on a computer screen for timing and spatial accuracy of hand motions during target-directed tasks (Chua & Elliott, 1993; Chua & Elliott, 1997; Miall, 1996; Miall & Reckess, 2002).

In contrast, several other studies have found similar results to the present study, suggesting that concurrent visual feedback offers little benefit for timing parameters in handwriting and tracking tasks (Brown & Donnenwirth, 1990), and may reduce timing accuracy (Majaranta, MacKenzie, Aula, & Räihä, 2006). For example, Ceux et al. (2003), who found that timing synchronization between hand motions and a moving target was worse when participants were permitted vision for matching the spatial and temporal properties of the stimulus. They remarked that their findings of negative effects of vision were counterintuitive, since vision was expected to form a stronger coupling between perception and action than a no-vision condition. They concluded that "it might very well be that the pursuit of a meticulous point-to-point spatio-temporal matching hampers the fluent execution of the cyclical movement" (p.105). Likewise, Marquardt, Gentz, & Mai, (1999) found concurrent visual feedback to function as a distractor in a handwriting task, particularly for tasks where the writers have achieved automaticity. These results are consistent with the actions of the participants in the present research, who overwhelmingly preferred to match the spatial properties of the target loops over the timing parameters of the target movements. Several researchers have proposed that visual feedback of spatial accuracy in tracking tasks produces competing demands upon limited cognitive resources, resulting in increased timing errors (Brown & Bennett, 2002; Brown & Donnenwirth, 1990).

It is not possible to assess the effects of concurrent visual feedback in the present study without understanding how the participants interpreted and used the visual trace of their recorded pen motions on the screen, which led to several surprises. For example, comments made by participants during non-systematic dialogues after the conclusion of their experimental sessions revealed that a considerable number of them did not recognize the visual trace as a representation of their own movements. Although differences in results from the Feedback:Present and Feedback:Absent conditions indicated that the participants were influenced by the visual trace, many participants did not perceive the visual trace as a record of their own movements. Instead, these participants stated that they thought that the visual trace was a distractor put on the screen by the experimenter while their hand motions drove the tempo of the target. Thus, the perception of their movements in relation to the moving target was distorted. The strength of this misperception was quite robust: several participants vigorously denied that the visual trace represented their own hand motions and could not be convinced otherwise. These participants reported that they viewed the visual presentation of their recorded traces as distractors, unrelated to their motions, least until the target completed its trajectory across the sample template. Thus, in a significant proportion of cases, participants may not have used the visual feedback from their hand motions as anticipated.

According to Betty Tuller, (personal communication, May 19, 2006), Professor of Complex Systems and Brain Sciences at Florida Atlantic University, misperceptions such as these are not uncommon. She reported that participants frequently inform researchers that they believe their movements to be driving the tempo of an oscillating target, even when visual feedback shows that their responses have lagged or led the stimulus. The perceived disconnection between an actor's movement and its visual representation is paradoxical, and can be assumed to influence the participants' use of information from visual feedback (Volcic, van Rheede, Postma, & Kappers, 2008). There are several possibilities to explain the ways in which participants used information from the visual trace in this study. One prospect is that the visual representation of the pen motions on the screen distracted participants from the tracking task, as the participants remarked. Another possibility is that the presentation of the visual trace with the moving target produced a dual task involving concurrent self-monitoring with pursuit loop tracking (Brown, & Bennett, 2002; Nielsen, 2006). In either case, visual feedback may have contributed to cognitive stress, resulting in slower moving times, and increased timing errors.

6.4.2. Intrasubject Variability at Landmarks across the Samples

Within each experimental condition, individual participants exhibited low variability at corresponding landmarks from trial to trial where peak-to-peak timing intervals were identical (i.e., in the TargetRate:Constant conditions and in the TargetRate:Changing conditions before and after the timing perturbation). Figure 69 in Chapter 5 shows that the points representing timing errors for one participant were close enough that they appeared to be superimposed on each other. The low trial-to-trial intrasubject variability in timing errors within the same experimental conditions was unexpected, since variability over repeated instances of a movement is one of the most common findings in motor control literature (Newell, Deutsch, Sosnoff, & Mayer-Kress, 2006). In contrast to these general observations on motor variability, aspects of motor consistency have been found in several studies of hand motions. For example, Viviani, Campadelli, and Mounoud (1987) found intertrial consistency in trajectories and timing errors in a smooth pursuit task. Athènes, Sallagoïty, Zanone, and Albaret, (2003; 2004) observed consistencies in specific shapes and eccentricities (slopes) of ellipses drawn by participants, which they hypothesized to suggest the presence of an attractor state emerging from characteristics of the actor-task system. Stelmach, Mullins, and Teulings (1984) found invariant timing features in a handwriting study and concluded that a well-specified motor program produced timing regularities. Roerdink et al. (2008) found motor consistency at anchor points in curved trajectories, commonly found where hand motions change direction, and attributed them to the influence of attractor effects at anchor points. Note that in the present study, no data were taken at any other points than the sample peaks, so no claim is made about consistency at any other location on the trajectory.

Several factors may have contributed to the timing regularities seen in these data. For example, participants were instructed to match their timing to the motion of the target while drawing the loops in the samples. Thus, the tempo of the moving target constrained the natural movement times of the participants, resulting in smaller timing differences between participants than would otherwise be observed if participants drew freely. Another possible contributor to the observed consistency is past practice with loop drawing. Although the tasks in the present study may have been novel in some aspects, past practice with similar tasks may have influenced their performance in the study, reducing or eliminating the variability occurring during motor learning. Previous studies found that timing errors may be significantly reduced in as few as three writing trials (Latash, 1999) or one hour of practice in a tracking task (Mosier, Scheidt, Acosta, & Mussa-Ivaldi, 2005). Thus, if the task was not sufficiently novel to the participants, they may have applied highly-rehearsed hand motions to the task. Since most participants did not adhere to the target tempo, they may have replicated movements at a well-rehearsed pace.

Another final possibility is that participants may have interpreted the task as too difficult to master in a short set of trials, and developed a compromise strategy to cope with competing spatial and timing demands of the task, as has been observed previously (Balasubramaniam, Wing, & Daffertshofer, 2004; Biberstine, Zelaznik, Kennedy, & Whetter, 2005). Kinsella-Shaw (personal communication, July 15, 2009) suggested that a unique self-organized solution may have emerged from the dynamics of the agent-task system, unfolding coincident with the task goal, without being enslaved to either the task requirements or the physical features of the effectors. Despite the fact that this explanation is far from conclusive, it is consistent with a growing body of literature that has reported that stable patterns emerge from the interaction between a rhythmic moving stimulus, task demands, and the coordinated activities of the mover (Kudo, Park, Kay, & Turvey, 2006; Roerdink et al. 2008) when the task exceeds the capability of the mover to achieve competing goals.

Although the low intrasubject variability in movement timing from trial to trial is an intriguing finding, it is not possible to assign scientific significance to this observation. A statistical analysis of the intertrial differences is beyond the scope of this study, and no normative data for comparison exist. Most important, the study was not designed to test theory regarding timing variability or consistency, and therefore cannot provide evidence in support or against any hypothesis that might explain these findings. Thus, further study on timing consistency at anchor points is warranted.

6.4.3. Drift (Temporal Lag) as a Specific Instance of Timing Error

In the present study, participants fell into idiosyncratic rates of rhythmic periodicity, lagging behind the visual stimulus even though they were instructed to keep their pen point in the leading dot of the moving target. Participants received no rewards for coupling their movements with the movement of the target, and no behavioral penalties for failing to do so. The results showed that a consistent pattern of progressive lag in timing (i.e., drift) occurred in most samples (98.6%), beginning at the first peak-to-peak segment and increasing with each subsequent cycle. As stated previously, timing errors were greater when participants obtained visual concurrent feedback on their performance (i.e., in the Feedback:Present conditions) than when the recorded trace of their movements was not provided (i.e., in the Feedback:Absent conditions).

Temporal drift is a well-documented finding in movement literature, cited as early as Stetson (1905) and Dunlap and Wells (1910). Stetson (1905) studied timekeeping with a baton, and found that timekeepers systematically increased the intervals between beats unless constrained by an external stimulus. Dunlap and Wells (1910) coined the term “drift” to describe increasing lag in a cycling activity. In tracking tasks, a lagging response to a temporal rhythm is typical, according to Viviani, Campadelli, and Mounoud (1987) and Miall and Reckess (1996). Lag is a consistent finding in pursuit tasks involving a cursor on a computer screen (MacKenzie & Ware, 1993), and progressively increasingly lags have been found when individuals are required to track more than one moving object (see Howard & Holcombe, 2008, for a review). Pellacchia and Turvey (2001) observed drift in a bimanual task, describing it in the language of dynamical systems and proposing that the operation of attractors in rhythmic activities causes participants to drift to preferred states. Approaching the issue of drift from the perspective of physics, Yates (1982) observed that the second law of thermodynamics produces gradual slowing in any motor task unless cognitive or physical energy is injected to maintain a steady rate. Thus, even though stable temporal relationships are expected in skilled tracking activities, lag remains a predictable outcome over time.

In reviewing the literature on lag in motor skills, Collier and Ogden (2004) revealed that most researchers have viewed the phenomenon of drift as motor error. They reported that various researchers manage the effects of drift in sequential data with a variety of methods, such as (a) studying very short intervals where little, if any lag would occur; (b) discarding data that exhibited lag; or (c) applying de-trending procedures to remove the lag between target and participant response. Collier and Ogden (2004) and Yu, Russell, and Sternad (2003), have questioned these strategies, however, and have suggested instead that movements whose timing and frequency gradually drifts from a timed stimulus may be an important area of study.

The systematic drift, or lag, seen in the data in the present study poses interesting questions about the perceptual-motor solutions to the tracking task used by participants. Once participants fell behind the target, it is possible that the cognitive efforts by participants to match the shape of the target negatively influenced prospective control processes (Allen, McGeorge, Pearson & Milne, 2004; Tombu & Seiffert, 2008). MacKenzie et al. (1993) determined that movers whose hand motions lagged behind a visual target were poorer at anticipating the rate of the target frequency, disrupting their prospective anticipation and prospective control of their motor responses to the movement of the visual stimulus, leading to increasingly greater lags as the participant traced across the sample. As the distance between target and hand motions expanded, participants may have found anticipatory control processes increasingly difficult to use, and fell into more familiar patterns.

6.5. Limitations of the Present Study

The purpose of this research was to test the role of concurrent visual feedback in tracking tasks relevant to handwriting and drawing performance. In brief, the results of the present study are consistent with findings of previous handwriting and tracking research, but the present study fails to provide significant support for those studies because of limitations in the test procedures. A major weakness of the experimental method used in this work is that it appears to lack ecological validity. Schmidt and Lee (2005) warned against the blindly accepting a common assumption that tasks performed in laboratory settings with feedback are fundamentally similar to ecologically relevant tasks in natural environments: "…using tasks so simple and artificial …may have little to tell us about the ways in which the rich and varied sources of inherent feedback work in natural settings" (p. 367-368). Therefore, the tasks in the present study may not effectively approximate the real-life handwriting tasks that the study was intended to illuminate. The following section describes several limitations posed by the instrumentation and methodology, which may limit the validity of this work.

6.5.1. Sample Size

The present study is limited by the size and characteristics of the study sample, which consisted of 23 female and 12 male students at the University of Minnesota, all of whom were right-handed. Although several left-handed students provided samples, there were too few left-handed students to include their data in the present study. Greater numbers of male and female participants, including left-handed participants, are necessary to generalize the findings to the college-aged population. Participants from other age groups are necessary to extend these findings to the broader population of healthy writers. An ideal sample would include children, adolescents, and older adults who are trained in a variety of writing systems (left-to-written Latin-based alphabets, right-to-left Semitic script, and pictographic writing systems, for example).

6.5.2. Use of Displaced Feedback

One important aspect of the present study was that the visual feedback provided to participants in eight (50%) of their 16 trials was displaced from the hand onto a monitor placed vertically at eye level. The use of a non-inking pen for drawing and the presentation of the recorded trace of their movement in real time on the screen were used to direct the participants' gaze onto the LCD monitor. Although a number of researchers have performed experimental studies of writing and drawing with a non-inking pen while observing the output of the hand motions in real time on a computer monitor (Ketcham, Dounskaia, & Stelmach, 2004; Nijhof, 2003; Reina & Schwartz, 2001; Overvelde & Hultsijn, 2009; Seidler-Dobrin & Stelmach, 1998; Teulings & Thomassen, 1979), all such experiments raise questions of ecological validity.

The present study faces the same issues, and, in particular, the methods in this study differ from an ecologically representative measure of handwriting in two important aspects:

(1) The feedback was a trace emanating directly from the pen but was displayed away from the pen and the hand that produced it.

(2) The present task involved writing on a horizontal surface while receiving feedback on a vertical surface.

With this arrangement, participants tracked an object presented vertically at eye level with hand motions made on a horizontal surface, requiring the participants to track the moving target by hand while observing the target and their response in different planes. In natural handwriting situations, individuals monitor their pen movements by observing the recorded trace in the same plane as the hand producing the trace. As a result, the method of displacing feedback in the Feedback:Present conditions in the present study raises questions about the naturalness of the participants' hand motions and their use of visual information emanating from their hand motions. Volcic et al. (2008) noted that effect of the separation between the hand and the visual feedback is not yet fully understood, and warned that information pertinent to effective perceptual-motor coupling might be lost when the hand motions and sources of information are separated. Thus, when feedback is displaced on the screen, a loss of information may ensue, particularly when participants are inexperienced with the task.

Anecdotal and research information suggests that such loss of information may not significantly degrade performance, however. Modern environments offer many tasks that require accurate perception and use of displaced feedback of hand motions, resulting in numerous studies conducted over the past fifty years that demonstrate rapid adaptation to displaced feedback (e.g., Held, Efstathiou, & Greene, 1966; Held, & Schlank, 1959). Driving, for example, involves a complex set of hand-to-task mapping procedures in which drivers rely on online visual information to move a vertically oriented steering wheel in order to keep the car operating within the narrow boundaries of a horizontal roadway (Gibson & Crooks, 1938; Land & Lee, 1994). Other examples include the use of a computer mouse (Elliott, Binsted, & Heath, 1999), a touchpad on a laptop (Dillen, Farris, & Meehan, 2005), a joystick (Miall & Reckess, 2002), video games, and Wii systems (Deutsch, Borbely, Filler, Huhn, & Guarrera-Bowlby, 2008). Latash (1999) found that displaced and distorted feedback posed few problems for effective performance since participants adjusted quickly to the mode of feedback. In his study, participants required only three practice sessions to write upside down and backwards well enough that their handwriting appeared normal on an inverted mirror. Even infants who have only recently acquired the skill of reaching quickly understand how to use a joystick to control a distant toy or generate self-motion on an infant-scaled robot (Galloway, Rhu, & Agrawal, 2008). Tasks such as these require the actor to both detect and control the motion in one plane while observing a visual representation of that motion displaced some distance away from the hand, occasionally on a different plane. Of particular relevance to the present study, Knoblich and Prinze (2001) have shown that self-generated motion on a horizontal plane can be effectively controlled when the visual record of movements is displaced onto a vertical computer screen. They found that participants who observed their horizontal tracking motions as point light displays shown on a vertical monitor were able to distinguish self-generated movements from other movements displayed on the screen.

Despite these examples, the question about the relevance to displaced feedback to natural writing and drawing tasks is pertinent to this study. In future studies, using instrumentation that would allow participants to write and track the traces of their hand motions more naturally in a horizontal plane would advance the understanding of the effects of concurrent visual feedback in hand motion tasks.

6.5.3. Dissimilarities between Tasks in the Present Study and Natural Handwriting

Most cursive writing includes meaningful content expressed by loops and other continuous and discrete shapes arranged into sequences of varying lengths. In the present study, the connected loops were designed from the cursive letter "e," using curves that are characteristic of letters in American cursive writing forms. However, the present task lacked semantic content and did not include variations in shape and length of writing units, reducing the generalizability of the experimental task to normal handwriting. Since the task in the present study did not replicate natural writing demands, it was not possible to determine whether it was representative of natural handwriting processes. Given these considerations, the present findings have limited applicability to natural handwriting tasks in typical environments.

6.5.4. The Tempo of the Experimental Target

In the present study, the use of the animated target was a device used to force study participants to adopt a uniform writing speed. The selected speed was expected to be obtained from judgments made by pilot participants (n = 30) as a speed that would enable most participants to draw at a natural pace while meeting the spatial requirements of the task. Determining an optimum target speed became much more difficult than anticipated because a majority of the pilot participants de-emphasized the timing parameters of the task to trace over the stationary template of twelve loops as carefully as possible, sometimes slowing their normal speeds by as much as 90%. Several of the faster participants admitted later that they thought the true purpose of the experiment was to see how quickly they could draw the loops, and so they ignored the spatial requirements of the task. As a result, the selected speed was not ideal for the task. Most participants in the experiment commented that they found the target speed too fast for concurrently achieving the timing and spatial goals of the task. In future studies, it would be useful to investigate tracking speeds more thoroughly by manipulating the speed of the target. This approach may help to determine critical speeds where participants abandon the effort to maintain both temporal and spatial coupling, and choose one movement accuracy goal, or alternatively, choose a compromise between the two timing and spatial accuracy demands.

6.6. Directions for Future Research

Replication is the test of any scientific endeavor, even when studying familiar actions. Since hand motions involved in handwriting are performed many times every day, there are many situations where changes in target-directed hand motions, including handwriting, drawing, and tracking, may illuminate perceptual, motor, and cognitive processes. The examples below are a small sample of research projects that could be performed.

6.6.1. Methodological Changes

6.6.1.1. Improve Ecological Validity

The present study yielded intriguing responses that warrant further investigation, but the issues of ecological validity must be first addressed. The system of displacing the visual representation of pen motions onto a vertical monitor should be replaced by a method that more closely approximates normal writing methods. One alternative approach would be the use of a tablet PC in which hand motions and the visual trace of the drawing are viewed in the same plane, as they are in the natural writing environment. To reduce the artificiality of the drawing task, a broader range of shapes that are foundational to modern writing methods should be obtained from handwriting and drawing samples, including continuous strokes such as circles and arches, and discrete strokes such as dots and lines. The effect of directionality should be assessed by examining samples written naturally in left-to-right , right-to-left, and vertical pictograph writing systems. Further, the dynamics and kinematics of movement when producing shapes without semantic content should be compared to movements used when forming meaningful words. These approaches would enhance understanding of cognitive, perceptual, and biomechanical aspects of natural writing.

6.6.1.2. Use Zero Crossing Velocity Measures to Segment Curves and Pressure

Réjean Plamondon, Professor of Electrical Engineering at École Polytechnique de Montréal has suggested that the sample be segmented by kinematic rather than spatial variables in the present study, which used a peak-to-peak segmentation process (personal communication, September 15, 2009). Specifically, he recommended that samples be segmented at zero crossings where decelerations and accelerations occur, suggesting that the zero-crossing method would reveal additional dynamics of pursuit loop-drawing performance and determine whether the wiggle observed in the samples is related to patterns of increasing and decreasing acceleration.

Pressure patterns may provide another mechanism for segmenting writing samples since they assist dynamic control in handwriting, and have been found to integrate and reveal underlying sensory, biomechanical, psychological, and cognitive processes (Kao, Hong, & Wah, 1986). Thus, observations of varying pressures under different writing conditions may advance understanding of the nature of multilevel systems operating during an instance of writing or drawing. Pressure was not evaluated in the present study, but appears to be a fertile area for further investigation in tracking, tracing, and drawing tasks (Kao et al., 1986).

6.6.1.3. Increase the Number and Types of Trials

In the present study, participants performed four trials of each experimental condition, randomized over the testing session. Intrasubject variability from trial to trial was minimal in many cases, perhaps because there were too few trials for motor learning to occur. In novel temporal coupling tasks, participants converge toward the timing of the external stimulus (Schmidt, Richardson, Arsenault, & Galantucci, 2007) with increased practice, but that phenomenon was not observed in the present study. Increasing the number of trials may provide insight into the processes of motor learning, and identify factors that foster temporal convergence toward the rhythm of the external stimulus.

Adding trials with different repetition schedules (for example, blocked vs. random practice) and varying target speeds may further clarify processes involved in motor learning. Using blocked (i.e., massed) practice rather than random practice and extending the number of practice sessions over several days may be useful in determining the amount and kinds of practice needed for learning cyclical drawing tasks and, perhaps, handwriting character shapes (allographs), graphemes (Rey, Ziegler, & Jacobs, 2000), and words.

6.6.1.4. Increase the Number and Types of Stimuli used to Entraining the Response to the Target

Manipulating the types of stimuli used to foster temporal convergence toward the target may also yield different patterns in the participants' responses. The present study allowed continuous feedback of the visual trace of the pen motions. However, discrete signals may provide different effects. For example, an auditory signal has been more effective in fostering entrainment with the tempo of an external stimulus than a visual signal (Chen, Repp, & Patel, 2002; Repp & Penel, 2002; 2004). The current research could be extended to include auditory cues, and evaluated for their effect on fostering convergence toward the target tempo. Another type of stimulus used for enhancing temporal coupling is a discrete visual signal, such as highlighting organizing points on the trajectory (Grant & Spivey, 2003). Thus, it may be useful to explore lag by comparing results when coupling is signaled by a continuous stream of visual information to discrete stimuli such as auditory tones or a discrete visual signals.

6.6.1.5. Investigate Perception of the Mode of Visual Feedback Used in this Study

Additional tools should be incorporated into the study design to assess how participants perceived the visual trace representing their hand movements. As stated previously, participants experienced an unanticipated response to the representation of the recorded traces of their pen movements on the screen, since many did not recognize the traces as their own. Thus, it is not clear whether they monitored them concurrently while tracking the moving object with their eyes, looked intermittently at them, or attempted to ignore them altogether. Eye movement technology to measure gaze and eye movements or fMRI analyses may help determine how the moving targets and other visual information were perceived (Christensen, Lundbye-Jensen, Petersen, Geertsen, Paulson, & Nielsen, 2006) and used to to guide movements during the experimental session.

6.6.2. Extending Research Questions

6.6.2.1. Examine the Effects of Concurrent Visual Feedback and Cognitive Load on Wiggle

The possibility that increases in wiggle are related to an increase in cognitive load with visual feedback could be tested by determining whether patterns of wiggles across writing samples are associated with changes in other movements known to be associated with cognitive load. For example, studies on vision and hand motions have shown that task-specific synergies develop during tasks involving eye-hand coordination

(Pelz, Hayhoe, & Loeber, 2001; Stuyven, Van der Goten, Vandierendonck, Claeys, & Crevits, 2000) but that the synergy may be disrupted by increasing cognitive load (Frens & Erkelens, 1991). These studies have identified a synergy between saccades and hand movements, but could be extended to test how saccades and wiggles are related, determining whether visual operations and movement patterns underlying wiggles are linked. Further tests could be performed to determine the relative contributions of visual operations, movement hesitations, muscular stiffness, velocity, pressure, or neuromotor noise (Van Galen & Van Huygevoort, 2000; Van Galen & Van Gemmert, 1996, 1997) to the incidence of wiggle, and assess the degree to which those factors can be manipulated by changes in perceptual or cognitive stress.

6.6.2.2. Investigate Timing Consistency Across Trials

The timing deviations at corresponding peaks in the same experimental conditions were smaller in the present study than expected, although consistency over repetitions has been observed in previous studies (Viviani, Campadelli, & Mounoud, 1987). However, the scientific significance of such apparent consistency is not known. In previous studies, temporal deviations from trial to trial as small as 2 milliseconds have been considered typical (Arend Van Gemmert, personal communication, May 12, 2009). The analysis of the trial-to-trial data requires complex statistical analysis that should be addressed in future studies. Such a study would require increasing the number of participants to establish normative measures for typical timing deviations in repeated trials in healthy populations. Information gained from normative data would be useful in assessing normal and clinical samples for transient, progressive, or chronic changes in perception and coordination which would assist in evaluating functional skill levels.

6.6.2.3. Evaluate the Influence of Attractor States in Handwriting

Another promising area of inquiry involves investigating the function of attractor states emerging from features of the task-agent system observed in cyclical hand motions in writing. Since the research described in this dissertation does not establish an ecological relevance for the experimental tasks used in this study, references to attractor states in participant responses are exploratory, but the low intrasubject variability at corresponding landmarks across trials warrants further study. In future studies with improved ecological validity, the operation of attractor states may be explored by investigating patterns of movement stability (Athènes et al., 2004; Beek, 1986; Buekers, Bogaerts, Swinnen, & Helsen, 2000), resistance to perturbations (Strogatz, 1994), and rapid reestablishment of previous patterns after minor perturbations, as demonstrated by short relaxation times (Kelso & Schöner, 1988). According to Jean-Michel Albaret, Professor, University Paul Sabatier, such a study would require testing a wider range of elliptical shapes and eccentricities than the ones used in the present study to ascertain whether loops translated across a page settle in elliptical shapes associated with attractors observed in past studies (personal communication, September 15, 2009).

6.6.2.4. Include Participants from Diverse Populations

As described previously when discussing limitations of the study, the information from the present research may not apply to the broader population. Future research should include participants from many age groups, including children, adolescents, and young, middle-aged, and older adults. Additional insights into the handwriting motions may be obtained by comparing healthy groups to clinical populations, including patients with chronic diseases, cognitive deficits, or undergoing pharmacologic or medical treatment, particularly to assess for negative effects on fine motor control. Extending this research to patients receiving chemotherapy may help assess neuropathy and cognitive effects of chemotherapy which are difficult for patients to assess and report. Other groups that may be included in future studies are people engaged in deception (for example, forgers). Extending these research protocols to studies comparing truthful and deceptive individuals (both sociopathic and non-sociopathic) may illustrate changes in fine motor control that reflect perceptual and cognitive processes involved in forgery and writing false statements.

6.6.2.5. Motivation for Future Research

The proposed studies described in the preceding section on study limitations show that investigations of fine motor control, particularly hand motions involved in writing and drawing are a rich area for future research. Investigations into hand motions in normal and clinical situations hold promise for increasing our knowledge of basic psychological, cognitive, neurological, and biomechanical processes and their variations under a host of conditions. Studies of hand motions involved in drawing and writing may thus provide abundant information useful for understanding our unique abilities for perceiving, responding, and adapting within complex social and physical environments.

REFERENCES

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3,* 111-150.

Adams, J. A., & Goetz, E. T. (1973). Feedback and practice as variables in error detection and correction. *Journal of Motor Behavior*, 5, 217-224.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* AC-19(6): 716–723.

Allen, R., McGeorge, P., Pearson, D., & Milne, A. (2004). Attention and expertise in multiple target tracking, *Applied Cognitive Psychology, 18*(3), 337–347.

Armstrong, D.F., & Wilcox, S.E. (2007). *The gestural origin of language*. New York: Oxford University Press.

Arbib, M.A. (2000). The Mirror System, Imitation, and the Evolution of Language. In C. Nehaniv & K. Dautenhahn, (Eds.), *Imitation in animals and artifacts* (pp. 229-280), Cambridge, MA: The MIT Press.

Asicioglu, F., & Turan, N. (2003). Handwriting changes under the effect of alcohol. *Forensic Science International, 132 (3),* 201-210.

Athènes, S., Sallagoïty, I., Zanone, P.G., & Albaret, J.M. (2003). Universal features of handwriting: Towards a non-linear model. *Proceedings of the 11th Conference of the International Graphonomics Society (IGS2003),* 46-49. Nijmegen, Holland: International Graphonomics Society.

Athènes, S., Sallagoïty, I., Zanone, P.G., & Albaret, J.M. (2004). Evaluating the coordination dynamics of handwriting. *Human Movement Science, 23*(5)*,* 621-641.

Baayen, R.H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Balasubramaniam, R. (2006). Trajectory formation in timed repetitive movements. In M.L. Latash & F. Lestienne (Eds.), *Motor control and learning* (pp. 47-54), NY: Springer.

Balasubramaniam, R., Wing. A.M., & Daffertshofer, A. (2004). Keeping with the beat: movement trajectories contribute to movement timing. *Experimental Brain Research, 159*(1), 129-134.

Bardy**,** B.G., & Laurent, M. (1998). How is body orientation controlled during somersaulting? *Journal of Experimental Psychology: Human Perception and Performance, 24*(3)*,* 963-977.

Barnes, D. E., Yaffe, K., Satariano, W.A., & Tager, I. B. (2003). A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *Journal of the American Geriatrics Society, 51*(4), 459.

Baron, D. E. (2009). *A better pencil.* New York: Oxford.

Barrientos, F.A. (2002). *Controlling expressive avatar gesture*. University of California, Berkeley, Ph.D. Dissertation, EECS Department, Computer Science Division.

Bavelas, J., Gerwing, J.J., Sutton, C.L., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, *58*(2), 495-520.

Beek, P. J. (1986). Timing and phase locking in cascade juggling. *Ecological Psychology*, *1*(1), 55-96.

Beek, P.J., Peper, C. E., & Daffertshofer, A. (2002). Modeling rhythmic interlimb coordination: Beyond the Haken–Kelso–Bunz model*, Brain and Cognition, 48*(1), 149-165.

Berninger, V.W., Abbott, R.D., Augsburger, A., & Garcia, N. (2009). Comparison of pen and keyboard transcription modes in children with and without learning disabilities. *Learning Disability Quarterly, 32*(3), 123-141.

Bernstein, N. (1967). *The co-ordination and regulation of movements*. Oxford: Pergamon Press.

Berryhill, M., Kveraga, K., Boucher, L., & Hughes, H.C. (2004). Smooth pursuit under stimulus–response uncertainty. *Cognitive Brain Research*, *19*(1), 100-102.

Beversdorf, D. Q., Anderson, J. M., Manning, S. E., Anderson, S. L., Nordgren, R. E., Felopulos, G. J., & Bauman, M. L. Brief report: macrographia in high-functioning adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 31*(1), 97-101.

Bingham, G.P. (2004). A perceptually driven dynamical model of bimanual rhythmic movements (and phase entrainment). *Ecological Psychology, 16,* 45–53.

Black, D., Found, B., & Rogers, D. (2003). The frequency of the occurrence of handwriting performance features used to predict whether questioned signatures are simulated. *Journal of Forensic Document Examination, 15,* 17-28.

Bongaardt, R., & Meijer, O.G. (2000). Bernstein's theory of movement behavior: Historical development and contemporary relevance. *Journal of Motor Behavior*, *32*(1), 57-71.

Bosga, J., Meulenbroek, R.G.J., & Rosenbaum, D. (2005). Deliberate control of continuous motor performance. *Journal of Motor Behavior, 37*(6), 437–446.

Box-Steffensmeier, J.M., & Jones, B.S. (2004). *Event history modeling: A guide for social scientists*. New York: Cambridge University Press.

Brault, J-J., & Plamondon, R.(1993). Segmenting handwritten signatures at their perceptually important points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(9), 953-957.

Breslow, N.E., & Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association, 88* (421), 9-25.

Bressler, S.L., & Kelso, J.A.S. (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences, 5*(1)*,* 26-36.

Brown, S. W., & Bennett, E. D. (2002). The role of practice and automaticity in temporal and nontemporal dual-task performance. *Psychological Research*, *66*(1), 80-89**.**

Brown, T. L., & Donnenwirth, E.E. (1990). Interaction of attentional and motor control processes in handwriting. *American Journal of Psychology*, *103*(4), 471-486.

Brown, D. G., Goovaerts, P., Burnicki, A., & Li, M.Y. (2002). Stochastic simulation of land-cover change using geostatistics and generalized additive models. *Photogrammetric Engineering and Remote Sensing, 68*(10):1051-1061.

Brown, J. S., McDonald, J. L., Brown, T. L., & Carr, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human Perception and Performance, 14*(1), 45-59.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). The neural basis of human dance. *Cerebral Cortex, 16*(8), 1157-1167.

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Buekers, M. J., Bogaerts, H. P., Swinnen, S. P., & Helsen, W. F. (2000). The synchronization of human arm movements to external events. *Neuroscience Letters, 290*(3), 181–184.

Bulacu, M., & Schomaker, L. (2007). *Automatic handwriting identification on medieval documents.* Paper presented at the 14th International Conference on Image Analysis and Processing (ICIAP 2007), Modena, Italy, pp.279-284. Retrieved May 12, 2008 from http://doi.ieeecomputersociety.org/10.1109/ICIAP.2007.33.

Burgess-Limerick, R., Abernethy, B., & Neal, R.J. (1993). Relative phase quantifies interjoint coordination. *Journal of Biomechanics, 26*(1), 91–94.

Burr, V. (2002). Judging gender from samples of adult handwriting: Accuracy and use of cues. *Journal of Social Psychology, 142(6),* 691-700.

Burton, A. W., Pick, H. L., Holmes, C., & Teulings, H. L. (1990). The independence of horizontal and vertical dimensions in handwriting with and without vision. *Acta Psychologica, 75*(3)*,* 201-212.

Burton, C., Strauss, E., Hultsch, D., Moll, A., & Hunter, M. (2006). Intraindividual variability as a marker of neurological dysfunction: a comparison of Alzheimer's disease and Parkinson's disease. *Journal of Clinical & Experimental Neuropsychology, 28*(1), 67-83.

Caligiuri, M.P., Teulings, H.L., Dean, C.E., Niculescu, A.B., & Lohr, J.B. (2009). Handwriting movement analyses for monitoring drug-induced motor side effects in schizophrenia patients treated with risperidone. *Human Movement Science, 28*(5), 633-642.

Caligiuri, M. P., Teulings, H. L., Filoteo, J. V., Song, D., & Lohr, J. B. (2006) Quantitative measurement of handwriting in the assessment of drug-induced parkinsonism. *Human Movement Science, 25*(4-5), 510-22.

Caporossi, G., Alamargot, D., & Chesnet, D. (2004). Using the computer to study the dynamics of handwriting processes. In S. Einoshin & A. Setsuo, (Eds.)*, Lecture Notes in Computer Science*, *Volume 3245* (pp. 242-254). Berlin: Springer.

Carroll, R.J. (2002). "Variances are not always nuisance parameters." Joint Statistical Meeting, Fisher Lecture 2002, New York City, 14 Sept. 2002. Retrieved April 9, 2009, from http://www.stat.tamu.edu~carroll/talks.php.

Carson, R.G . (1996). Neuromuscular-skeletal constraints upon the dynamics of perception-action coupling. *Experimental Brain Research, 110 (1)*, 99-110.

Case, P., Tuller, B., & Kelso, J.A.S. (2003, November 17). The dynamics of learning to hear new speech sounds. *SpeechPathology.Com.* Retrieved from http: //www.speechpathology.com/articles/arc_disp.asp?article_id=50&catid=560

Castiello, U., & Stelmach, G. E. (1993). Generalized representation of handwriting: Evidence of effector independence. *Acta Psychologica, 82*(1-3)*,* 53–68.

Ceux, T., Buekers, M.J., & Montagne, G. (2003). The effects of enhanced visual feedback on human synchronization. *Neuroscience Letters, 349*(2), 103-106.

Chatterjee, S., & Hadi, A. S. (2006). *Regression analysis by example*. Hoboken, NJ: Wiley-Interscience.

Chen H.C., Cha, S.H. Chee, Y.M., & Tappert, C. (2003). The detection of forged handwriting using a fractal number estimate of wrinkliness. *Proceedings of the 11th Conference of the International Graphonomics Society (IGS2003),* 312-315. Nijmegen, Holland: International Graphonomics Society.

Chen, Y., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science, 21*(4), 515 - 532.

Chen, A., Weng, X., Yuan, J., Lei, X., Qiu, J., Yao, D., & Li, H. (2008). The temporal features of self-referential processing evoked by Chinese handwriting. *Journal of Cognitive Neuroscience, 20*(5), 816-827.

Christensen, M.S., Lundbye-Jensen, J., Petersen, N., Geertsen, S. S., Paulson, O. B., & Nielsen, J. B. (2006). Watching your foot move--An fMRI study of visuomotor interactions during foot movement. *Cerebral Cortex, 17* (8), 1906-1917.

Chua, R., & Elliott, D. (1993). Visual regulation of manual aiming. *Human Movement Science, 12*(4), 365-401.

Chua, R., & Elliott, D. (1997). Visual control of target-directed movements. *Behavioral and Brain Sciences, 20*, 304-306.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences,* 3(9), 345-351.

Clayton, M., Sager, R., & Will, U. (2005). In time with the music: the concept of entrainment and its significance for ethnomusicology*. European Meetings in Ethnomusicology, 11*, 3-142.

Cohen, P., Cohen, J., West, S.G., & Aiken, L.S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences.* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum.

Collier, G. L., & Ogden, R. T. (2004). Adding drift to the decomposition of simple isochronous rhythmic tapping: An extension of the Wing–Kristofferson model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 853–872.

Connelly, V. (2009, September). *Language and writing development: how language problems can constrain the development of writing in children*. Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Consiglio, W., Driscoll, P., Witte, M., & Berg, W. P. (2003). Effect of cellular telephone conversations and other potential interference on reaction time in breaking response. *Accident analysis and prevention, 35*(4)*,* 495-500.

Cook, R. D., & Weisberg, S. (1999). *Applied regression including computing and graphics*. NY: John Wiley.

Coolidge, F. L. (2006). *Statistics: A gentle introduction*. Thousand Oaks, CA: Sage Publications.

Copple M. (2003). Gesture and speech: André Leroi-Gourhan's theory of the co-evolution of manual and intellectual activities. *Gesture*, *3*(1), 47-94.

Corballis, M. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences, 26(2)*, 199–260.

Coull, B.A., Schwartz, J., & Wand, M.P.W. (2001). Respiratory health and air pollution: additive mixed model analyses. *Biostatistics***,** 2(3), 337-349.

Crawley. M. (2005). *Statistics: An introduction using R.* West Sussex, England: John Wiley.

Crawley. M. (2007). *The R Book*. West Sussex, England: John Wiley.

Dalgaard, P. (2002). *Introductory Statistics with R.* NY: Springer.

Dayton, C. M. (2003). Model comparisons using information measures. *Journal of Modern Applied Statistical Methods, 2*(2), 281-292.

Dean, M. P., & Gilmore, R. (2008). Myoclonus affecting speech reduced by concurrent handwriting. *Movement Disorders*, *23*(2), 306-312.

De Jong, W. P., Hulstijn, W., Kosterman, B.J.M., & Smits-Engelsman, B.C.M. (1996). OASIS software and its application in experimental handwriting research. In M.L. Simner, C. J. Leedham, & A. J. W. M. Thomassen (Eds.), *Handwriting and drawing research: Basic and applied issues* (pp. 429-440). Amsterdam: IOS Press.

de Leeuw, J., & Meijer, E. (Eds.). (2008). *Handbook of multilevel analysis.* NY:Springer.

Denier van der Gon, J.J., & Thuring, J.P. (1965). The guiding of human writing movements. *Biological Cybernetics, 2*(4), 145-148.

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and lifestyle survey. *Psychology and Aging 21*(1): 62-73.

de'Sperati C., & Viviani P. (1997). The relationship between curvature and velocity in two-dimensional smooth pursuit eye movements. *Journal of Neuroscience, 17*(10), 3932-3945.

Deutsch, J. E., Borbely, M., Filler, J., Huhn, K., & Guarrera-Bowlby, P. (2008). Use of a low-cost, commercially available gaming console (wii) for rehabilitation of an adolescent with cerebral palsy. *Physical Therapy, 88*(10), 1196-1207.

Deutsch, K. M., & Newell, K. M. (2005). Noise, variability, and the development of children’s perceptual-motor skills. *Developmental Review, 25*(2), 155-180.

Devore, J. L., & Peck, R. (2000). *Statistics: The exploration and analysis of data.* Pacific Grove, CA: Duxbury.

Dhawan, B. N., Bapat, S.K., & Saxena, V.C. (1969). Effect of four centrally acting drugs on handwriting. *The Japanese Journal of Pharmacology, 19*(1), 63-67.

Dick, D., Found, B., & Rogers, D. (2000). The forensic detection of deceptive behavior using handwriting movements. *Journal of Forensic Document Examination, 13*, 15-24.

Dillen, H., Farris, J. S.., & Meehan, J, W. (2005). Kinematic analysis of cursor trajectories controlled with a touchpad*. International Journal of Human-Computer Interaction, 19*(2), 223 – 239.

Dixon, R.A., Kurzen, D.F., & Friesen, I.C. (1993.) Handwriting performance in older adults: Age, familiarity and practice effects. *Psychology and Aging, 8*(3), 360–370.

Djebbari, R., du Montcel, S. T., Sangla, S., Vidal, J. S., Gallouedec, G., & Vidailhet, M. (2004). Factors predicting improvement in motor disability in writer's cramp treated with botulinum toxin. *Journal of Neurology, Neurosurgery and Psychiatry, 75*(12), 1688–1691.

Dose, M., Gruber, C., Grunz, A., Hook, C., Kempf, J., Scharfenberg, G., & Sick, B.(2007, April). Towards an automated analysis of neuroleptics' impact on human hand motor skills. *Proceedings of the IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology, 2007* (pp. 494-501).

Doumas, M., & Wing, A.M. (2007). Timing and trajectory in rhythm production. *Journal of Experimental Psychology: Human Perception and Performance, 33* (2), 442–455.

Dounskaia, N. (2007). Kinematic invariants during cyclical arm movements. *Biological Cybernetics, 96*(1), 147–163.

Dounskaia, N., Van Gemmert, A.W.A., & Stelmach, G.E. (2000). Interjoint coordination during handwriting-like movements. *Experimental Brain Research, 135*(1), 127–140.

Downey, J.E. (1910). Review of control processes in modified hand - writing: an experimental study. *Psychological Bulletin, 7(5)*, 165-167.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. NY: Wiley.

Dunlap, K., & Wells, G. R. (1910). Reactions to rhythmic stimuli, with attempt to synchronize. *Psychological Review*, *17,* 399–416.

Edwards, A.L. (1984). *Experimental design in psychological research.* NY: HarperCollins.

Elble, R. J. (2003). Characteristics of physical tremor in young and elderly adults. *Clinical Neurophysiology, 114*(4) 624-635.

Elble, R. J. (2004). Tremor in ostensibly normal elderly people. *Movement Disorders, 13*(3), 457-464.

Elliott, D., Binsted, G., & Heath, M. (1999). The control of goal-directed limb movements: Correcting errors in the trajectory. *Human Movement Science, 18*(2-3), 121-136.

Elliott, D., Garson, R.G., Goodman, D., & Chua, R. (1991). Discrete vs. continuous visual control of manual aiming. *Human Movement Science*, *10*(4), 393-418.

Evidente, V.G. (2000). Understanding essential tremor: Differential diagnosis and options for treatment. *Postgraduate Medicine, 108* (5), 138-140, 143-146.

Fairclough, S. F., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology, 56*, 171-184.

Faraway, J. J. (2006). *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models*. Boca Raton, FL: Chapman & Hall/CRC.

Farrer, C., Bouchereau, M., Jeannerod, M., & Franck, N. (2008). Effect of distorted visual feedback on the sense of agency. *Behavioural Neurology, 19*, 53–57.

Feldman, R. (2007) Parent-infant synchrony: biological foundations and developmental outcomes. *Current Directions in Psychological Science, 16* (6), 340–345.

Field, A. (2005). *Discovering statistics using SPSS*. London: Sage.

Fink, P., Foo, P. Jirsa, V.K., & Kelso, J.A.S., (2000). Local and global stabilization of coordination by sensory information. *Experimental Brain Research, 134*(1), 9-20.

Finkel, E. J., Campbell, W. K., Brunell, A.B., Dalton, A. N., Scarbeck, S.J., & Chartrand, T. L. (2006). High-maintenance interaction: inefficient social coordination impairs self-regulation. *Journal of Personality and Social Psychology, 91*(3), 456-475.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381-391. (Reprinted in *Journal of Experimental Psychology: General*, *121*(3), 262-269, 1992.)

Florey, K. B. (2009). *Script and Scribble: The rise and fall of handwriting.* Brooklyn, NY: Melville House.

Foley, R.G. (1987). Characteristics of synchronous sequential signatures. *Journal of Forensic Sciences, 32*(1), 121-129.

Fowler, C. A., Richardson, M. J., Marsh, K. L., & Shockley, K. (2008). Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. Jirsa (Eds). *Coordination: Neural, behavioral and social dynamics*. NY: Springer.

Fox, J. (2002). *Linear mixed models: Appendix to an R and S-PLUS companion to applied regression*. Downloaded May 24, 2009 from http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-mixed-models.pdf.

Fox, J. (2008). *Applied regression analysis and generalized linear models*. Thousand Oaks, CA: Sage.

French, J. L., & Wand, M. P. (2004). Generalized additive models for cancer mapping with incomplete covariates. *Biostatistics*, 5(2), 177-191.

Frens, M.A., & Erkelens, C.J. (1991) Coordination of hand movements and saccades: evidence for a common and a separate pathway. *Experimental Brain Research, 85*(3), 682–690.

Frøkjær-Jenson, B. (1980). *The sciences of deaf signing.* Copenhagen: København University Press.

Galloway, J.C., Rhu, J., & Agrawal, S. (2008). Babies driving robots: Self-generated mobility in very young infants. *Intelligent Service Robotics, 1*(2).

Galwey, N. (2006). *An introduction to mixed modelling: beyond regression and analysis of variance.* Hoboken, NJ: Wiley.

Ganz, R. E., Ehrenstein, W. H., & Cavonius, C. R. (1996). Dynamic complexity of visuo-motor coordination: an extension of Bernstein's conception of the degrees-of-freedom problem. *Biological Cybernetics*, *75*(5), 381-387.

Gelman, A. (2005). Analysis of variance—why it is more important than ever. *Annals of Statistics, 33*(1), 1–53.

Getchell, N. (2006). Age and task-related differences in timing stability, consistency, and natural frequency of children's rhythmic, motor coordination. *Developmental Psychobiology, 48*, 675-685.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

Gibson, J.J. (1969). A suggested classification of the types and subtypes of graphic action. Unpublished manuscript. Retrieved February 20, 2007 from http://www.trincoll.edu/depts/ecopsyc/perils/folder4/classification.html.

Gibson, J.J. (1972). A theory of direct visual perception. In J. Royce & W. Rozenboom (Eds.). *The psychology of knowing*. New York: Gordon & Breach.

Gibson, J.J., & Crooks, L.E. (1938) A theoretical field-analysis of automobile-driving. *American Journal of Psychology, 51*, 453–471.

Gibbs, R. W. (2005). *Embodiment and cognitive science.* NY: Cambridge University Press.

Gill, S. P. (2007) Entrainment and musicality in the human system interface. *AI & Society: Knowledge, Culture, and Communication*, *21*(2), 567–605.

Giraudo, M.D., & Pailhous, J. (1999). Dynamic instability of visual images. *Journal of Experimental Psychology. Human Perception and Performance, 25*(6)*,* 1495-1516.

Glencross, D. J. (1980). Levels and strategies of response organization. In G. E. Stelmach & J. Requin (Eds.). *Tutorials in motor behavior*. Amsterdam: North-Holland.

Gnanadesikan, R. (1977). *Methods for statistical analysis of multivariate data*. NY: Wiley.

Goodnow, J.J., & Levine, R.A. (1973). The grammar of action: Sequence and syntax in children's copying. *Cognitive Psychology, 4*(1), 82–98.

Graham, S., & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review, 8*(1), 7-87.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving. *Psychological Science, 14* (5), 462-466.

Gross, L. (1975). Drug-induced handwriting changes: an empirical review. *Texas Reports on Biology and Medicine, 33* (3), 370-390.

Guadagnoli M.A., & and Kohl R.M. (2001) Knowledge of results for motor learning: relationship between error estimation and knowledge of results frequency. *Journal of Motor Behavior, 33*(2), 217–224.

Guadagnoli, M. A., Dornier, L. A., & Tandy, R. (1996). Optimal summary knowledge of results: The influence of task related experience and complexity. *Research Quarterly for Exercise and Sport, 67,* 239–248.

Guadagnoli, M.A., Leis, B., Van Gemmert, A.W., & Stelmach G.E. (2002). The relationship between knowledge of results and motor learning in Parkinsonian patients. *Parkinsonism & Related Disorders, 9*(2), 89-95.

Guigon, E., Baraduc, P., and Desmurget, M. (2008). Computational motor control: Feedback and accuracy. *European Journal of Neuroscience, 27* (4), 1003-1016.

Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling, 157,* 89-100.

Haddock, V. (2007, February 25). We shouldn't write off handwriting just yet. *San Francisco Chronicle*, p. E-1. Retrieved January 7, 2008 from http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2007/02/25/INGALO8UUB1.DTL#ixzz0TVpxBfRq.

Haken, H., Kelso, J.A.S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*(5), 347–356.

Hancock, P. A., & Newell, K. M. (1985). The movement speed-accuracy relationship in space-time. In H. Heuer, U. Kleinbeck, & K.-H. Schmidt (Eds.), *Motor behavior: Programming, control and acquisition* (pp. 153-188). New York: Springer-Verlag.

Harralson, H. H., Teulings, H.L., & Farley, B. G. (2008). Comparison of handwriting kinematics in movement disorders and forgery. *Journal of Forensic Document Examination, 19,* 41-52.

Hartley, L.R. (1981) Noise, attentional selectivity, serial reactions and the need for experimental power. *British Journal of Psychology, 72*(1), 101-107.

Hastie, T.J., & Tibshirani, R.J. (1990). *Generalized additive models.* Boca Raton: Chapman & Hall/CRC.

Hayes, W. N. (1996). Identifying sex from handwriting. *Perceptual and Motor Skills, 83* (3, Part 1), 791-800.

Heath, R.A. (2000). *Nonlinear dynamics.* Mahwah, NJ: Lawrence Erlbaum.

Held, R., Efstathiou, A., & Greene, M. (1966). Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology, 72*, 887-891.

Held, R., & Schlank, M. (1959). Adaptation to disarranged eye-hand coordination in the distance dimension. *American Journal of Psychology, 72*, 603-605.

Hick, W.E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quarterly Journal of Experimental Psychology (A)*, *1*, 36-51.

Hill, A.R., Adams, J.M., Parker, Jr., B. E., & Rochester, D. F. (1988). Short-term entrainment of ventilation to the walking cycle in man. *Journal of Applied Physiology, 65,* 570-578.

Hilton, O. (1969). Study of the influence of alcohol on handwriting. *Journal of Forensic Sciences, 14,* 309-16.

Hinder, M. R., Tresilian, J.R., Riek, S., & Carson, R. G. (2008). The contribution of visual feedback to visuomotor adaptation: How much and when? *Brain Research, 1197*, 123-134.

Hirsch, M.W., Jarvik, M.E., & Abramson, H.A. (1956) Lysergic acid diethylamide (LSD-25): XVIII. Effects of LSD-25 and six related drugs upon handwriting. *Journal of Psychology, 41*, 11-22.

Hirt, L. (1899). *The diseases of the nervous system. A text-book for physicians and students.* (A. Hoch & F.R. Smith, Trans.) New York: Appleton.

Hollerbach, J. M. (1981). An oscillation theory of handwriting. *Biological Cybernetics, 39*(2), 139-156.

Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research, 48*(9), 1164-1180.

Howell, D.C. (2008). *Mixed models for repeated (longitudinal) data*. Retrieved December 12, 2008 from http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data /Mixed%20Models%20for%20Repeated%20Measures.pdf.

Hox, J. J. (2002). *Multilevel analysis: techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hsieh, S. 2002. Task shifting in dual-task settings. *Perceptual and Motor Skills, 94*(2), 407.

Hung, S. W-S., Molnar, G. F., Ashby, P., Chen, R., Voon, V., & Lang, A. (2005). Electrophysiologic testing in psychogenic tremor: does it always help? In M. Hallett, C.R. Cloninger, S. Fahn, J.J. Jankovic & A.E. Lang (Eds.), *The psychogenic movement disorders: neurology and neuropsychiatry* (4th ed., pp. 334-335). Philadelphia: Lippincott, Williams & Wilkins.

Hurley, S. R., & Lee, T. D. (2006). The influence of augmented feedback and prior learning on the acquisition of a new bimanual coordination pattern. *Human Movement Science, 25*(3), 339-348.

Iglowstein, I., Jenni, O.G., Molinari, L., & Largo, R. H. (2003). Sleep duration from infancy to adolescence: Reference values and generational trends. *Pediatrics, 111*(2), 302-307.

Ihler, A.T., Fisher, J.W., & Willsky, A.S. (2001). Nonparametric estimators for online signature authentication. *Proceedings of the 2001 IEEE International Conference on Acoustics, Speech and Signal Processing*, *6* (pp. 3473—3476).

Ivry, R. B., & Richardson, T.C. (2002). Temporal control and coordination: the multiple timer model. *Brain and Cognition, 48*, 117–132.

Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.M., & Thiébaut, R. (2007). Robustness of the linear mixed model to misspecified error distribution. *Computational Statistics & Data Analysis, 51*(10), 5142-5154.

Jasper, I., Gordijn, M., Haüssler, A., Marquardt, C., & Hermsdörfer, J. (2009). Circadian rhythm in handwriting. *Journal of Sleep Research, 18*(2), 264-271.

Jennrich, R. I., & Schluchter, M.D. (1986) Unbalanced repeated-measures models with structured covariance matrices. *Biometrics, 42*(2), 805-820.

Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review 96*(3), 459-491.

Kagerer, F.L., Contreras-Vidal, J. L., Bo, J., & Clark**,** J.E. (2006). Abrupt, but not gradual visuomotor distortion facilitates adaptation in children with developmental coordination disorder. *Human Movement Science, 25* (4-5), 622-633*.*

Kaiser, M.K., Proffitt, D.R., Whelan, S.M., & Hecht, H. (1992). The influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18 (3)*, 669-690.

Kandel, S. (2009, September). *For a psycholinguistic approach of handwriting production.* Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Kandel, S., Álvarez, C. J., & Vallée, N. (2006). Syllables as processing units in handwriting production. *Journal of experimental psychology: Human perception and performance, 32*(1), 18-31.

Kandel, S., Orliaguet, J.P., & Boë L. J. (2000) Detecting anticipatory events in handwriting movements. *Perception, 29*(8), 953-64.

Kao, H.S.R., Hong, M.P., & Wah, L. P. (1986). Handwriting pressure: Effects of task complexity, control mode, and orthographic difference. In H.S.R. Kao, G. P. van Galen, & R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting* (pp. 47-66). Amsterdam: Elsevier.

Kao, H. S. R., Shek, D. T. L., & Lee, E. S. P. (1983). Control modes and task complexity in tracing and handwriting performance. *Acta Psychologica, 54*(1-3)*,* 69-77.

Kaplan, D., & Glass, L. (1995). *Understanding nonlinear dynamics*. NY: Springer-Verlag.

Karls, I., Maderlechner, G., Pflug, V., Baumann, S., Weigel, A., & Dengel, A. (May, 1993). Segmentation and recognition of cursive handwriting with improved structured lexica. *Proceedings of the Third International Workshop Frontiers in Handwriting Recognition (IWFHRIII)*, *Buffalo, NY,* 437-442.

Kawashima, R., Tajima, N., Yoshida, H., Okita, K., Sasaki, T., Schormann, T., Ogawa, A., Fukuda, H., & Zilles, K. (2000). The effect of verbal feedback on motor learning—A PET study. *NeuroImage, 12*(6), 698-706.

Kay, B. A., Saltzman, E.L., Kelso, J.A.S., & Schöner, G. (1987). Space-Time behavior of single and bimanual rhythmical movements: Data and limit cycle model. *Journal of Experimental Psychology: Human Perception and Performance, 13 (2)*, 178-192.

Keele, Luke J. (2006). *How to be smooth: Smoothing in political science*. Retrieved April 7, 2009 from http://www.polisci.ohio-state.edu/faculty/lkeele/smoothersver1.pdf.

Keele, Luke J. (2008). *Semiparametric regression for the social sciences.* NY:Wiley.

Keith, T.Z. (2006). *Multiple regression and beyond.* Boston: Pearson.

Kelly, S.D., Iverson, J.M., Terranova, J., Niego, J., Hopkins, M., & Goldsmith, L. (2002). Putting language back in the body: Speech and gesture on three time frames. *Developmental Neuropsychology, 22*(1), 323-349.

Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Kelso, J.A.S. (1997). Relative timing in brain and behavior: Some observations about the generalized motor program and self-organized coordination dynamics. *Human Movement Science, 16*(4), 453-460.

Kelso, J.A.S. (2003). Cognitive coordination dynamics. In W. Tschacher & J.P. Dauwalder (Eds.), *The dynamical systems approach to cognition: Concepts and empirical paradigms based on self-organization, embodiment and coordination dynamics* (pp. 45-71). Singapore: World Scientific.

Kelso, J.A.S., DelColle, J.D., & Schöner, G. (1990). Action perception as a pattern formation process. In M. Jeanerod (Ed.), *Attention and performance XIII* (pp. 139-169). Hillsdale, NJ: Erlbaum.

Kelso, J.A.S., & Schöner, G. (1988). Self-organization of coordinative movement patterns. *Human Movement Science, 7*(1), 27–46.

Kelso, J.A.S., Schöner, G., Scholz, J.P., & Haken, H. (1987). Phase-locked modes, phase transitions and component oscillators in coordinated biological motion, *Physica Scripta, 35 (1),* 79-87.

Keogh, J., Morrison, S., & Barrett, R. (2004). Augmented visual feedback increases finger tremor during postural pointing. *Experimental Brain Research*, *159*(4), 467-477.

Ketcham, C., Dounskaia, N., & Stelmach, G. (2006). The role of vision in the control of continuous multijoint movements. *Journal of Motor Behavior, 38*(1), 29–44.

Kim, W., & Linton, O. (2004). The live method for generalized additive volatility models. *Econometric theory*, *20*(6), 1094-1139.

Klapp, S. T. (1996). Reaction time analysis of central motor control. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 13-35). Champaign, IL: Human Kinetics.

Klein, R. (2001). Fully modern humans. In G. M. Feinman & T. D. Price (Eds.), *Archaeology at the millennium: A sourcebook*. NY: Kluwer Academic.

Knoblich, G., & Prinze, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology: Human Perception and Performance. 27*(2), 456-465.

Kording K.P., & Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427*, 244–247.

Kreft, I.G.G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Kudo, K., Park, H., Kay, B. A., & Turvey, M. T. (2006). Environmental coupling modulates the attractors of rhythmic coordination. *Journal of Experimental Psychology: Human Perception and Performance, 32*(3), 599-609.

Kugler, P. N., Kelso, J. A. S., & Turvey, M. T. (1980). On the concept of coordinative structures as dissipative structures: I. Theoretical lines of convergence. In G.E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 3-47). Amsterdam: North-Holland.

Kurgansky, A.V. (2008). Visuomotor synchronization: Analysis of the initiation and stable synchronization phases. *Human Physiology, 34*(3), 289-298.

Lacquaniti, F., Terzuolo, C.A., Viviani, P. (1983). The law relating kinematic and figural aspects of drawing movements. *Acta Psychologica, 54*(1-3), 115-130.

Lagarde, J., & Kelso, J. A. S. . (2006) Binding of movement, sound and touch: multimodal coordination dynamics. *Experimental Brain Research, 173*(4), 673-688.

Laird, N. M., & Ware, J.H. (1982). Random effects models for longitudinal data. *Biometrics, 38*(4), 936-974.

Lames, M. (2006). Modelling the interaction in game sports – relative phase and moving correlations. *Journal of Sports Science and Medicine, 5*, 556-560.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature, 369*(6483)**,** 742 – 744.

Lange, K., Tucha, O., Aschenbrenner, S., Gottwald, D., Eichhammer, P., Putzhammer, A., Sartor, H., & Klein, H. (2001). The effect of antidepressant medication on handwriting of patients with depression. In R.G.J. Meulenbroek, & B. Steenbergen (Eds.), *Proceedings of the 10th Biennial Conference of the International Graphonomics Society* (pp. 52-57). Nijmegen, Holland: International Graphonomics Society.

Lange, K.W., Tucha, O., Aschenbrenner, S., Mecklinger, L., Eichhammer, P., & Klein, H. (2001). Kinematic analysis of handwriting movements in patients with untreated schizophrenia. In R.G.J. Meulenbroeck & B. Steenbergen (Eds.), *Proceedings of the 10th Biennial Conference of the International Graphonomics Society* (pp. 183-187). Nijmegen, Holland: International Graphonomics Society.

Lange, K.W., Tucha, O., Mecklinger, L., & Paul, G.M. (2003). Antidepressants and graphomotor skills of patients with depression. *The Forensic Scientist, 6*, 10-15.

Lange, K.W., Tucha, L., Walitza, S., Gerlach, M., Linder, M., & Tucha, O. (2007). Interaction of attention and graphomotor functions in children with attention deficit hyperactivity disorder. *Journal of Neural Transmission (Suppl)*, 72, 249-259.

Lashley, K. S. (1930). Basic neural mechanisms in behavior. *Psychological Review, 37,* 1–24.

Latash, M. L. (1999). Mirror writing: learning, transfer, and implications for internal inverse models. *Journal of Motor Behavior, 31*(2), 107-110.

Leopold, J.L., & Ambler, A.L. (April, 1997). Keyboardless visual programming using voice, handwriting, and gesture. *1997 IEEE Symposium on Visual Languages (VL '97)*. Retrieved October 8, 2009 from http://doi.ieeecomputersociety.org /10.1109/VL.1997.626555.

Leroi-Gourhan, A. (1993). *Gesture and speech* (A.B. Berger, Trans.). Cambridge, MA: MIT.

Lewinson, T. S., & Zubin, J. (1942). *Handwriting analysis. A series of scales for evaluating the dynamic aspects of handwriting*. New York: King's Crown Press.

Liao, M., & Jagacinski, R.J. (2000). A dynamical systems approach to manual tracking performance. *Journal of Motor Behavior*, *32*(4), 361-378.

Lin, X., & Zhang, D. (1999). Inference in generalized additive mixed models by using smoothing splines. *Journal of the Royal Statistical Society Series B*, *61*(Part 2), 381-400.

Linderman, M., Lebedev, M., & Erlichman, J. (August 26, 2009). Recognition of handwriting from electromyography*. Public Library of Science ONE (PloS, August 26, 2009).* Retrieved October 10, 2009 from http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0006791.

Lindstrom, M. J., & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data, *Biometrics*, *46*(3), 673-687.

Littell, R, Milliken, G., Stroup, W., Wolfinger, R., & Schabenberger, O. (2006). *SAS for mixed methods* (2nd ed.). Cary, NC: SAS Publishing.

Liu, X., & Scheidt, R. A. (2008). Contributions of online visual feedback to the learning and generalization of novel finger coordination patterns. *Journal of Neurophysiology, 99*(5), 2546-2557.

Longford, N. T. (1994). *Random coefficient models*. NY: Oxford University Press-USA.

Longstaff, M.G., & Heath, R.A. (1997). Space-time invariance in adult handwriting. *Acta Psychologica, 97(2),* 201-214.

Longstaff, M.G., & Heath, R.A. (1999). A nonlinear analysis of the temporal characteristics of handwriting. *Human Movement Science, 18(4)*, 485-524.

Longstaff, M.G., & Heath, R.A. (2000). The influence of tremor on handwriting performance under conditions of low and intermediate physical stress. *The Journal of Forensic Document Examination, 13*, 25-44.

Longstaff, M.G., & Heath, R.A. (2003). The influence of motor system degradation on the control of handwriting movements: A dynamical systems analysis. *Human Movement Science*, *22*(1), 91-110.

Longstaff, M.G., & Heath, R.A. (2006). Spiral drawing performance as an indicator of fine motor function in people with multiple sclerosis. *Human Movement Science, 25*(4-5), 474-491.

Lopresti-Goodman, S. M., Richardson, M. J., Silva, P., & Schmidt, R. C. (2008). Period basin of entrainment for unintentional visual coordination. *Journal of Motor Behavior*, *40*(1), 3-10.

Louis, E. D., Wendt, K. J., & Ford, B. (2000). Senile tremor: What is the prevalence and severity of tremor in older adults? *Gerontology, 46,* 12–16.

Luke, D.A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage.

Luria, G., & Rosenblum, S. (2009). Comparing the handwriting behaviours of true and false writing with computerized handwriting measures. *Applied Cognitive Psychology.* Published online. doi: 10.1002/acp.1621.

Maarse, F.J., Van Galen, G.P., & Thomassen, A.J. (1989). Models for the generation of writing units in handwriting under variation of size, slant, and orientation. *Human Movement Science*, *8*(3), 271–288.

MacKenzie, I. S., & Ware, C. (1993). Lag as a determinant of human performance in interactive systems. *Proceedings of the ACM Conference on Human Factors in Computing Systems - INTERCHI '93*, 488-493. New York: ACM.

Majaranta, P., MacKenzie, I. S., Aula, A., & Räihä, K.-J. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society (UAIS)*, *5*, 199-208.

Marquardt, C., Gentz, W., & Mai, N. 1996. On the role of vision in skilled handwriting. In M.L. Simner, C.G. Leedham, & A.J.W.M. Thomassen (Eds.), *Handwriting and drawing research: Basic and applied issues* (pp. 87-97). Amsterdam: IOS Press.

Marquardt, C., Gentz, W., & Mai, N. (1999). Visual control of automated handwriting movements. *Experimental Brain Research*, 128 *(1-2)*, 224-228.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated-measures designs: some alternative approaches. *Child Development, 44*(3), 401-415.

McLennan, J.E., Nakano, K., Tyler, H.R., & Schwab R.S. (1972). Micrographia in Parkinson's disease. *Journal of the Neurological Sciences, 15,* 141–52.

McGarva, A.R., & Warner, R.M. (2003). Attraction and social coordination: mutual entrainment of vocal activity rhythms. *Journal of Psycholinguistic Research*, 32(3), 335-354.

Medwell, J., Strand, S., Wray, D. (2008). What should we assess in primary handwriting? In A. Barnett (Ed.). *Handwriting Today* (23-28). Oxford: National Handwriting Association.

Mergl R., Juckel G., Rihl J., Henkel, V., Karner, M., Tigges, P., Schröter, A., & Hegerl, U. (2004). Kinematical analysis of handwriting movements in depressed patients. *Acta Psychiatrica Scandinavica, 109*, 383-391.

Mergl, R., Tigges, P., Schroter, A., Moller, H.J., Hegerl, U. (1999). Digitized analysis of handwriting and drawing movements in healthy subjects: Methods, results and perspectives. *Journal of Neuroscience Methods, 90*(2): 157-69.

Merton, P.A. (1972). How we control the contraction of our muscles. *Scientific American, 226,* 30-37.

Meulenbroek, R.G.J., Thomassen, A.J.W.M., Rosebaum, D. A., Loukopoulos, L. D., & Vaughan, J. (1996). Adaptation of a reaching model to handwriting: How different effectors can produce the same written output, and other results. *Psychological Research*, *59*(1), 64-74.

Meulenbroek, R.G.J., Thomassen, A.J.W.M., Schillings, J.J., & Rosenbaum, D.A. (1996). Synergies and sequencing in copying L-shaped patterns. In: Simner, M.L., Leedham, C.G., Thomassen, A.J.W.M. (Eds.), *Handwriting and drawing research: Basic and applied issues* (pp. 41-55). Amsterdam: IOS Press.

Meulenbroek R.G.J., Thomassen A.J.W.M., van Lieshout P.H.H.M., & Swinnen S.P. (1998). The stability of pen–joint and interjoint coordination in loop writing. *Acta Psychologica, 100*(1-2), 55-70.

Meulenbroek, R. G. J., Van Galen, G. P., Hulstijn, M., Hulstijn, W., & Bloemsaat, G. (2005). Muscular co-contraction covaries with task load to control the flow of motion in fine motor tasks. *Biological Psychology, 68 (3),* 331-352.

Meulenbroek, R.G.J., & van Gemmert, A.W.A. (2003). Advances in the study of drawing and handwriting. *Human Movement Science, 22*(2), 131-135.

Meyer D.E., Smith J.E., & Wright C.E. (1982). Models for the speed and accuracy of aimed movements. *Psychological Review, 89*, 449-482.

Miall, R. C. (1996) Task-dependent changes in visual feedback control: A frequency analysis of human manual tracking. *Journal of Motor Behavior, 28*(2), 125–135.

Miall R. C., & Reckess G. Z. (2002). The cerebellum and the timing of coordinated eye and hand tracking. *Brain and Cognition, 48*(1), 212-226.

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

Montgomery, G. (2002). The ancient origins of sign handshapes. *Sign Language Studies, 2*(3), 322-334.

Morasso, P. (1986). Understanding cursive script as a trajectory formation paradigm. In H.S.R. Kao, G.P. van Galen, & R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting* (pp. 137-167). Amsterdam: Elsevier.

Morrens, M, Wezenberg, E. Verkes, R. J., Hulstijn, W., Ruigt, G., Sabbe, B., (2007). Psychomotor and memory effects of haloperidol, olanzapine, and paroxetine in healthy subjects after short-term administration. *Journal of Clinical Psychopharmacology, 27*(1), 15-21.

Mosier, K. M., Scheidt, R.A., Acosta, S., & Mussa-Ivaldi, F. A. (2005). Remapping hand movements in a novel geometrical environment. *Journal of Neurophysiology, 94,* 4362-4372.

Naito, H., Chang, W-D., & Shin, J. (2009, September). *Detecting the effect of alcohol consumption using on-line handwritten signature.* Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Nelder, J., & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General), 135,* 370–384.

Neter, J., Wasserman, W., & Kutner, M.H. (1990). *Applied linear regression models* (3rd ed.). Homewood, IL: Irwin.

Newell, K.M., & Carlton, M.J. (1987). Augmented information feedback and the acquisition of isometric tasks. *Journal of Motor Behavior, 19*(1), 4–12.

Newell, K.M. Carlton, L.G., & Seonjim, K. (1994). Time and space-time movement accuracy. *Human Performance, 7*(1),1-21.

Newell, K.M., Deutsch, K.M., Sosnoff, J.J., & Mayer-Kress, G. (2006). Variability in motor output as noise: A default and erroneous proposition? In K. Davids, S. Bennett, K.M. Newell (Eds). *Movement system variability* (pp. 3-24), Champaign: Human Kinetics.

Newell, K.M., Liu, Y.T., & Mayer-Kress, G., 2001. Time scales in motor learning and development. *Psychological Review, 108*(1), 57-82

Newell, K.M., & Van Emmerik, R.E.A. (1989). The acquisition of coordination: Preliminary analysis of learning to write. *Human Movement Science, 8* (1), 17-32.

Newell, K.M., & Slifkin, A. B. (1998). Nature of movement variability. In Jan Piek (Ed.), *Motor behavior and human skill: A multidisciplinary approach* (pp.143-160). Champaign: Human Kinetics.

NIST/SEMATECH. Quantile-Quantile Plot. *NIST/SEMATECH e-Handbook of Statistical Methods* (§1.3.3.24). Retrieved April 5, 2008 from http://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm.

Neisser, U. (1967). *Cognitive psychology*. Upper Saddle River, New Jersey: Prentice Hall.

Nijhof, E-J. (2003). On-line trajectory modifications of planar, goal-directed arm movements. *Human Movement Science, 22*(1), 13–36. doi:10.1016/S0167-9457(02)00140-9.

Niizeki, K., & Miyamoto, Y. (1998). Cardiolocomotor interactions during dynamic handgrip and knee extension exercises: Phase-locked synchronization and its physiological implications. *Advances in Experimental Medicine and Biology*, 450, 199-206.

Noble, M., Fitts, P.M., Warren, C. E. (1955). The frequency response of skilled subjects in a pursuit tracking task. *Journal of Experimental Psychology, 49*(4), 249-256.

NUI Group Authors (2009). *Multi-touch technologies*. Retrieved October 8, 2009, from http://nuicode.com/attachments/download/115/Multi-Touch_Technologies_ v1.01.pdf.

Numerical Algorithms Group Ltd. (2002). NAG Fortran Library Chapter Introduction: G10 – Smoothing in statistics. Retrieved January 17, 2009 from http://www.nag.com/numeric/fl/manual20/pdf/G10/g10_intro.pdf.

Oehlert, G.W. (2000). *A first course in design and analysis of experiments*. NY: Freeman.

Olive, T. (2004). Working memory in writing: Empirical evidences from the dual-task technique. *European Psychologist, 9*(1), 32–42

O'Reilly, C., Plamondon, R., Clément, B., Mathieu, P., & Lebrun, L.H. (2009). *Movement analysis for stroke susceptibility assessment*. Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Orliaguet JP, Kandel S, Boë L.J. (1997). Visual perception of motor anticipation in cursive handwriting: Influence of spatial and movement information on the prediction of forthcoming letters. *Perception* 26(7):905-12.

Ouillier, O., Bardy, B., & Bootsma, R.J. (2001). Intrinsic and informational coupling in stance. In G.A. Burton & R.C. Schmidt (Eds.), *Studies in perception and action VI* (pp. 169-172). Hillsdale, NJ: Erlbaum.

Overvelde, A.J.A.A.M & Hulstijn, W. (2009). *A comparison of three methods of learning to write a new letterform.* Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Park, J-H., Shea, C. H., & Wright, D. L. (2000). Reduced frequency concurrent and terminal feedback: A test of the guidance hypothesis. *Journal of Motor Behavior, 32*(3), 287-296.

Patel, A.D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception, 24*, 99-104.

Patterson H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika 58*(3), 545-554.

Pellecchia, G.L., Shockley, K., & Turvey, M.T. (2005). Concurrent cognitive task modulates coordination dynamics. *Cognitive Science: A Multidisciplinary Journal, 29*(4), 531-557.

Pellecchia, G., & Turvey, M. T. (2001). Cognitive activity shifts the attractors of bimanual rhythmic coordination. *Journal of Motor Behavior*, *33*(1), 9-15.

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research, 139*(3), pp. 266–277.

Perl, W. R. (1955). On the psychodiagnostic value of handwriting analysis. A*merican Journal of Psychiatry, 111*(8), 595-602.

Phelps, J., & Stempel, L. (1987). Handwriting: Evolution and evaluation. *Annals of dyslexia, 37*(1), 228-239.

Phillips, J.G., Bradshaw, J.L., Chiu, E., & Bradshaw J.A. (1994). Characteristics of handwriting of patients with Huntington's disease. *Movement Disorders, 9*(5), 521-30.

Phillips, J.G., Ogeil, R. P., & Rogers, D. (2008). Alcohol consumption, dependence and the spatial relationships of handwriting. *Journal of Forensic Document Examination*, *19*, 5-28.

Phillips-Silver, J., & Trainor, L.J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105 (3)*, 533-546.

Pinheiro, J. C., & Bates, D. M. (2004). *Mixed effects models in S and S-Plus*. Springer-Verlag: New York.

Pinheiro, J. C., Bates, D. M., DebRoy, S., Sarkar, D., & the R Core Team (December 30, 2008). The nlme package. Retrieved January 2, 2009 from http://cran.r-project.org/web/packages/nlme/nlme.pdf.

Plamondon, R., & Alimi, A.M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences, 20*(2), 279-349.

Plamondon, R., Alimi, A., Yergeau, P., & Leclerc, F. (1993). Modelling velocity profiles of rapid movements: A comparative study. *Biological Cybernetics*, *69*(2), 119-128.

Plamondon, R., & Guerfali, W. (1996). Why handwriting segmentation can be misleading? *Proceedings of the 13th International Conference on Pattern Recognition, Vol.D* (pp. 396-400), Vienna, Austria.

Plamondon, R., & Srihari, S. (2000). On-Line and off-line handwriting recognition: a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1),* p. 63-84.

Port, R., Tajima, K., & Cummins, F. (1996). Self-entrainment in animal behavior and human speech. In M. Gasser (Ed.), *Online Proceedings of the 1996 Midwest Artificial Intelligence and Cognitive Science Conference*. Retrieved October 14, 2007 from http://www.cs.indiana.edu/event/maics96/Proceedings/Port/port.html.

Pothier, M. (2007, May 6). In digital age, more t's are crossed poorly. *Boston Globe.* Retrieved May 9, 2007 from http://www.boston.com/business/globe/articles/2007/05/06/in_digital_age_more_ts_are_crossed_poorly/.

Preyer, W. (1895). *Zur Psychologie des Schreibens*. Leipzig: Verlag von Leopold Voss.

Proteau, L.,Marteniuk, R. G. (1993). Static visual information and the learning and control of a manual aiming movement. *Human Movement Science*, *12*(5), 515-536.

Proteau, L.,Marteniuk, R. G., Girouard, Y., & Dugas, C. (1987). On the type of information used to control and learn an aiming movement after moderate and extensive training. *Human Movement Science, 6*(2), 181-199.

Proteau, L., Tremblay, L., & DeJaeger, D. (1998). Practice does not diminish the role of visual information in online control of a precision walking task: Support for the specificity of practice hypothesis. *Journal of Motor Behavior, 30*(2), 143–150.

Proteau, L., Cournoyer, J. (1990). Vision of the stylus in a manual aiming task: The effects of practice. *The Quarterly Journal of Experimental Psychology, 42A*(4), 811 – 828.

Qu, Tong (2004). *Dynamic signature verification system design using stroke based feature extraction algorithm*. Unpublished master's dissertation. Retrieved November 11, 2007, from http://www.sce.carleton.ca/faculty/adler /publications/2004/tongqu-MASc-thesis-2004.pdf.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing Versions 2.8,1 and 2.5. Vienna, Austria: R Foundation for Statistical Computing. Downloaded December, 22, 2008 from http://www.R-project.org.

Raethjen, J., Pawlas, F., Lindemann, M., Wenzelburger, R., & Deuschl, G.(2000). Determinants of physiologic tremor in a large normal population. *Clinical Neurophysiology, 111*(10), 1825-1837.

Raibert, M.H. (1977). *Motor control and learning by the state space model.* (Tech. Rep. No. 439). Cambridge, MA: MIT Press.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Ravizza, S. (2003). Movement and lexical access: Do noniconic gestures aid in retrieval? *Psychonomic Bulletin & Review, 10*(3), 610–615.

Reina, G.A., & Schwartz, A.B. (2003). Eye–hand coupling during closed-loop drawing: Evidence of shared motor planning? *Human Movement Science, 22*(2), 137-152.

Renze, J. Outlier. From *MathWorld*--A Wolfram Web Resource, created by Eric W. Weisstein. http://mathworld.wolfram.com/Outlier.html.

Repp, B.H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance, 27(3)*, 600-621.

Repp, B. H. (2002). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance. 28*(2), 410-430.

Repp, B.H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review, 12*(6), 969-992.

Repp, B.H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, *68*, 252-270.

Rey, A., Ziegler, J.C., & Jacobs, A.M. (2000). Graphemes are perceptual reading units. *Cognition, 75*(1), B1-B12.

Richardson M. J., & Flash T. (2002). Comparing smooth arm movements with the two-thirds power law and the related segmented-control hypothesis. *Journal of Neuroscience, 22,* 8201–8211.

Richardson, M. J., Schmidt, R. C., & Kay, B. A. (2007). Distinguishing the noise and attractor strength of coordinated limb movements using recurrence analysis. *Biological Cybernetics, 96*(1), 59-78.

Richardson, M. J., Lopresti-Goodman, S., Mancini, M., Kay, B. A., & Schmidt, R. C. (2008). Comparing the attractor strength of intra- and interpersonal interlimb coordination using cross recurrence analysis. *Neuroscience Letters, 438,* 340-345.

Richardson, M.J., Marsh, K.L., & Schmidt, R.C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 31 (1)*, 62-79.

Richardson, M.J., Marsh, K.L. Isenhower, R.W., Goodman, J. R. L., & Schmidt, R.C. (2007) Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science, 26*(6)*,* 867-891.

Rijntjes, M., Dettmers, D., Büchel, C., Kiebel, S., Frackowiak, R.S.J., & Weiller, C. (1999). A blueprint for movement: Functional and anatomical representations in the human motor system. *The Journal of Neuroscience, 19 (18)*, 8043-8048.

Robertson, S.D., Zelaznik, H. N., Lantero, D.A., Bojczyk, K.G., Spencer, R. M., Doffin, J. G., Schneidt, T. (1999). Correlations for timing consistency among tapping and drawing tasks: Evidence against a single timing process for motor control. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1316-1330.

Roerdink, M., Peper, C. E., & Beek, P. J. (2005). Effects of correct and transformed visual feedback on rhythmic visuo-motor tracking: Tracking performance and visual search behavior. *Human Movement Science, 24*(3), 379-402.

Roerdink, M., Ophoff, E. D., Peper, C. E., & Beek, P. J. (2008). Visual and musculoskeletal underpinnings of anchoring in rhythmic visuo-motor tracking. *Experimental Brain Research, 184*(2), 143–156.

Roitman, A., Johnson, M.T.V., & Ebner, T.J. (2001). Kinematic analysis of manual tracking: Interception and pursuit. *Society for Neuroscience Abstracts,* 27, 940-947.

Rosenblum, S., Chevion, D., & Weiss, P.L.T. (2006). Using data visualization and signal processing to characterize the handwriting process. *Pediatric Rehabilitation, 9*(4), 404-417.

Ruppert, D. Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression*. Cambridge: Cambridge University Press.

Sabbe, S.B.G., Hulstijn, W., Pier, M. B. P. I., & Zitman, F. G. (2000). Psychomotor retardation in treatment of depression with SSRI's. *European Psychiatry, 15*(2), 218.

Sage, I., Zesiger, P., & Garitte, C. (2009, September). *The links between motor development and the development of handwriting among eight- to twelve-year-olds.* Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Sallagoïty, I., Athènes, S., Zanone, P.-G., & Albaret, J.-M. (2003). Alterations of graphic production under various constraints follow principled modifications in underlying patterns dynamics. *Proceedings of the 11th Conference of the International Graphonomics Society (IGS2003),* 50-54. Nijmegen, Holland: International Graphonomics Society.

Sallagoïty, I., Athènes, S., Zanone, P.-G., & Albaret, J.-M. (2002). Probing the coordination dynamics of handwriting. In Toussaint, L & Boulinguez, P. (Eds), *Proceedings of the International Congress of Movement, Attention, and Perception*, p. 137.

Sallagoïty, I., Athènes, S., Zanone, P.-G., & Albaret, J.-M. (2004). Stability of coordination patterns in handwriting: Effects of speed and hand. *Motor Control, 8*(4), 405-421.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science, 19*(4), 499-526.

Sanguineti, V., Frisone, F., Bruni, S., & Morasso, P. (1998). Can non-linear muscle dynamics explain the smoothness of handwriting movements? *Acta Psychologica, 100 (1)*, 217-227.

Saura Ramos, P. A. (1999). Beltran, A. (Ed.). *The Cave of Altamira*. NY: Harry N. Abrams.

Schaal, S., Mohajerian, P. Ijspeert, A. (2007). Dynamics systems vs. optimal control – a unifying view. *Progress in Brain Research, 165,* 425-445.

Schmandt-Besserat, D. (1996). *How writing came about.* Austin, TX: University of Texas Press.

Schmandt-Besserat, D. (2007). *When writing met art.* Austin, TX: University of Texas Press.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*(4)*,* 225–260.

Schmidt, R. A. (1982). *Motor control and learning.* Champaign, IL: Human Kinetics.

Schmidt, R.A., & Lee, T.D. (2005). *Motor control and learning: a behavioral emphasis* (4th ed.) Champaign: Human Kinetics.

Schmidt, R.A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: implications for training and simulation. *Human Factors, 39*(4), 509-25.

Schmidt, R.C., Richardson, M.J., Arsenault, C., & Galantucci, B. (2007). Visual tracking and entrainment to an environmental rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4)*,* 860-870.

Schmidt, R.C., & Turvey, M.T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics, 70*(4), 369–376.

Schmuckler, M.A. (2001). What is ecological validity? A dimensional analysis. *Infancy, 2*(4), 419 – 436.

Schomaker, L. (2008). Writer identification and verification. (2008). In N. Ratha & V. Govindaraju, (Eds.), *Advances in biometrics: Sensors, systems and algorithms* (pp. 247-264). London: Springer Verlag.

Schomaker, L., & Teulings, H.L. (1992). Stroke- versus character-based recognition of on-line connective cursive script. In J. C. Simon & S. Impedovo (Eds.), *From pixels to features III* (pp. 313-325), Amsterdam: North-Holland.

Schöner, G. (1991). Dynamic theory of action-perception patterns: the "moving room" paradigm. *Biological Cybernetics, 64*(6), 455-462.

Schwarz, G., (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2):461-464.

Sciacca, E., Langlois-Peter, M-B., Gilhodes, J-C., Margot, P., & Velay, J-L. (2009). The range of handwriting variability under different writing conditions. *Journal of Forensic Document Examination, 19,* 5-13.

Searle, L.V., & Taylor, F.T. (1948). Studies of tracking behavior. I. Rate and time characteristics of simple corrective movements. *Journal of Experimental Psychology, 38,* 615-631.

Searle, S. R., Casella, G., & McCulloch, C.E. (1992). *Variance components.* NY: Wiley.

Seidler-Dobrin, R. D., & Stelmach, G. E. (1998). Persistence in visual feedback control by the elderly. *Experimental Brain Research, 119*(4), 467-474.

Seidler, R. D., Noll, D.C., & Thiers, G. (2004). Feedforward and feedback processes in motor control. *NeuroImage*, *22*(4), 1775-1783.

Sennef, C., Van Riezen, H., De Jong, W. P., Hulstijn, W. (n.d.). *Manual to the OASIS-based automated test-battery for psychopharmacological research.* Nijmegen: Kikosoft.

Shanahan, T. (2008). Relations among oral language, reading, and writing development. In C.A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 171-183). NY: Guilford.

Sharpe, K., & Van Gelder, L. (2006). The study of finger flutings. *Cambridge Archaeological Journal, 16*(3): 281-295.

Shaw, R., & Schockley, K. (2003). An ecological science of the artificial? *Journal of the Learning Sciences, 12*(3), 427-435.

Shepherd, B. (2008, December). *Dealing with endogeneity.* (December, 2008). Paper presented at the ARTNeT Capacity Building Workshop for Trade Research, UNESCAP Trade and Investment Division, Bangkok, Thailand.

Shockley, K., Santana, M., & Fowler, C.A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29 (2)*, 326-332.

Simner, M. L., Leedham, C. G., Thomassen, A. J. W. M. (Eds.). (2006). *Handwriting and drawing research: Basic and applied issues.* Amsterdam: IOS Press.

Simonoff, J.S. (1996). *Smoothing methods in statistics*. New York: Springer.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence.* NY: Oxford.

Singer, Y., & Tishby, N. (1994). Dynamical encoding of cursive handwriting. *Biological Cybernetics, 71*(3), 227-237.

Sita, J.C., Rogers, D., & Found, B. (2003). A model using complexity classification, spatial score and line quality for forensic signature comparison. *Proceedings of the 11th Biennial Conference of the International Graphonomics Society*, 257-260. Nijmegen, Holland: International Graphonomics Society.

Slavin, M. J., Phillips, J. G., Bradshaw, J. L., Hall, K. A., & Presnell, I. (1999). Consistency of handwriting movements in dementia of the Alzheimer's type: A comparison with Huntington's and Parkinson's diseases. *Journal of the International Neuropsychological Society, 5*, 20-25.

Smits-Engelsman, B. C., Niemeijer, A. S., van Galen, G. P. (2001). Fine motor deficiencies in children diagnosed as DCD based on poor graphomotor ability. *Human Movement Science, 20*(1-2), 161-182.

Smits-Engelsman, B.C., Wilson, P., Westenberg, Y., & Duysens, J. (2003). Fine motor deficiencies in children with developmental coordination disorder and learning disabilities: an underlying open-loop motor deficit. *Human Movement Science, 22*(4-5), 495-513.

Smyth, M.M., & J. Silvers, 1987. Functions of vision in the control of handwriting. *Acta Psychologica, 65*(1), 47-67.

Snijders, T.A.B., & Bosker, R. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Staude, G., Dengler, R., & Wolf, W. (2002). The discontinuous nature of motor execution II. Merging discrete and rhythmic movements in a single-joint system – the phase entrainment effect. *Biological Cybernetics, 86*(6), 427–443.

Stelmach, G.E., & Diggles, V. A. (1982). Control theories in motor behavior. *Acta Psychologica, 50*(1), 83-105.

Stelmach, G.E., Mullins, P.A., & Teulings, H.L. (1984). Motor programming and temporal patterns in handwriting. *Annals of the New York Academy of Sciences, 423,* 144 – 157.

Sternad, D., Turvey, M.T., & Schmidt, R.C. (1992). Average phase difference theory and 1:1 phase entrainment in interlimb coordination. *Biological Cybernetics, 67*(3), 223-231.

Stetson, R. H. (1905). A motor theory of rhythm and discrete succession: II. *Psychological Bulletin, 12* (5), 293-350.

Stoffregen, T.A., Hove, P., Schmit, J., & Bardy, B.G. (2006). Voluntary and involuntary postural responses to imposed optic flow. *Motor Control, 10*(1), 24-33.

Streufert, S., Streufert, S.C., Denson A. L. (1983). Information load stress, risk taking, and physiological responsivity in a visual-motor task. *Journal of Applied Social Psychology*, *13*(2), 145 – 163.

Strogatz, S.H. (1994). *Nonlinear dynamics and chaos.* Cambridge, MA: Perseus.

Strogatz, S.H. (2003). *Sync: The emerging science of spontaneous order.* New York: Hyperion.

Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., & Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta Psychologica, 104*(1)*,* 69-85.

Swinnen S.P., Dounskaia, N., Walter, C.B., Serrien, D.J. (1997). Preferred and induced coordination modes during the acquisition of bimanual movements with a 2:1 frequency ratio. *Journal of Experimental Psychology: Human Perception and Performance, 23*(4), 1087-1110.

Swinnen S.P., Schmidt, R.A., Nicholson, D.E., & Shapiro D.C. (1990) Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 706-716.

Tamhane, A. C., & Dunlop, D. D. (2000). *Statistics and data analysis from elementary to intermediate.* Upper Saddle River, NJ: Prentice Hall

Teulings, H.L., Contreras-Vidal, J. L., Stelmach, G. E., & Adler, C. H. (2002). Adaptation of handwriting size under distorted visual feedback in patients with Parkinson's disease and elderly and young controls. *Journal of Neurology, Neurosurgery and Psychiatry, 72,* 315-324.

Teulings, H. L., & Schomaker, L. R. B. (1993). Invariant properties between stroke features in handwriting. *Acta Psychologica, 82*(1-3), 69–88.

Teulings, H. L., Thomassen, A. J. W. M., & van Galen, G. P. (1983). Preparation of partly precued movements: The size of movement units in handwriting. *Acta Psychologica, 54*(1-3), 165-177.

Teulings, H. H. M., & Thomassen, A. J. W. M. (1979). Computer-aided analysis of handwriting movements. *Visible Language, 13*, 218-231.

Thelen, E., Corbetta, D., Spencer, J.P. (1996). The development of reaching during the first year: The role of movement speed. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1059-1076.

Thelen, E., & Smith, L.B. (2006). Dynamic systems theories. In W. Damon & R.M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (6th ed., pp. 258 – 312). New York: Wiley.

Thomassen, A.J.W.M & Meulenbroek, R.G.J. (1998). Low-frequency periodicity in the coordination of progressive handwriting. *Acta Psychologica, 100*(1-2), 133-144.

Thornton, T. P. (1996). *Handwriting in America: A cultural history.* New Haven: Yale.

Tiplady, B., Baird, R., Lütcke, H., Drummond, G., & Wright, P. (2005). Effects of ethanol on kinaesthetic perception. *Journal of Psychopharmacology, 19,* 627-632.

Tlili, M., Mottet, D., Dupuy, M.-A., Pavis, B. (2004). Stability and phase locking in human soccer juggling. *Neuroscience Letters, 360*(1-2),45-48.

Todorov, E., Shadmehr, R., Bizzi, E. (1997). Augmented feedback presented in a virtual environment accelerates learning of a difficult motor task. *Journal of Motor Behavior, 29*(2), 147-158.

Tombu, M., & Seiffert, A.E. (2008). Attentional costs in multiple-object tracking. *Cognition 108*(1), 1-25.

Trainor, L.J. (2007). Do preferred beat rate and entrainment to the beat have a common origin in movement? *Empirical Musicology Review, 2*(1), 17-20.

Treffner, P., & Turvey, M. T. (1993) Resonance constraints on rhythmic movement. *Journal of Experimental Psychology: Human Perception & Performance, 19*, 1221-1237.

Tucha, O., & Lange, K.W. (2004). Effects of nicotine chewing gum on a real-life motor task: A kinematic analysis of handwriting movements in smokers and non-smokers. *Psychopharmacology, 173*, 49-56.

Tucha, L., Lange, K.W., Stasik, D., Walitza, S., & Tucha, O.. (2008). Disturbances of handwriting in children with attention deficit hyperactivity disorder. *Handwriting, 7,* 29-36.

Tucha, L., Tucha, O., Laufkoetter, R., Walitza, S., Klein, H.E., & Lange, K.W.. (2008). Neuropsychological assessment of attention in adults with different subtypes of attention-deficit/hyperactivity disorder. *Journal of Neural Transmission, 115*, 269-278.

Tucha, O., Laufkötter, R., Mecklinger, L., Klein, H., Lange, K. (2001). Handwriting of adult patients with attention deficit hyperactivity disorder. In R.G.J. Meulenbroek, & B. Steenbergen (Eds.), *Proceedings of the 10th Biennial Conference of the International Graphonomics Society* (*IGS 2001*), (pp. 58-62). Nijmegen, Holland: International Graphonomics Society.

Tucha, O., Stasik, D., Mecklinger, L., & Lange, K.W. (2006). The effect of caffeine on handwriting movements in skilled writers. *Human Movement Science, 25*(4-5), 523-535.

Tuller, B., Turvey, M. T., & Fitch, H. L. (1982). The Bernstein perspective II: The concept of muscle linkage or coordinative structure. In J. A. Scott Kelso (Ed.), *Human motor behavior: An Introduction* (pp. 239-252). Hillsdale NJ: Lawrence Erlbaum.

Turvey, M.T. (1977). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 211-265). Hillsdale, NJ: Erlbaum.

Turvey, Michael T. (1990) Coordination. *American Psychology, 45,* 938-953.

Turvey, M. T., & Carello, C. (1981). Cognition: The view from ecological realism. *Cognition, 10* (1-3), 313-321.

Turvey, M. T., & Carello, C. (1986). The ecological approach to perceiving-acting: A pictorial essay. *Acta Psychologica, 63*(2)*,* 133-155.

Turvey, M. T., Shaw, R., & Mace, W. M. (1978). Issues in the theory of action: Degrees of freedom, coordinative structures and coalitions. In J. Requin (Ed.), *Attention and Performance VII* (pp. 557 – 595). Hillsdale, NJ: Lawrence Erlbaum.

Ueno, K., Furukawa, K., & Bain, M. (2000). Motor skill as dynamic constraint satisfaction. *Linking Electronic Articles in Computer and Information Science, 5* (36). Retrieved January 17, 2008 from http://www.ep.liu.se/ea/cis/2000/036.

Ulmanová, O., Homann, C.N., Ulman, R., Jech, R., Čapek, V., Klempíř, J., & Růžička, E. (2007). Tremor magnitude: A single index to assess writing and drawing in essential tremor. *Parkinsonism & Related Disorders, 13*(4), 250-253.

Underwood, A. J. (1997). *Experiments in ecology.* NY: Cambridge.

Ünlü, A., Brause, R., & Krakow, K. (2006). Handwriting analysis for diagnosis and prognosis of Parkinson's disease. In N.Maglaveras, I.Chouvarda, V. Koutkias, R.Brause (Eds.), *Proceedings of the International Symposium of Biological and Medical Data Analysis, 4345* (pp. 441-450). Heidelberg: Springer Verlag.

Van Den Heuvel, C.E., van Galen, G.P., Teulings, H.L., & van Gemmert, A.W.A. (1998). Axial pen force increases with processing demands in handwriting. *Acta Psychologica, 100*(1-2), 145-159.

Van Dijk**,** H. , Mulder, T., & Hermens, H. J. (2007). The effects of age and content of augmented feedback on learning an isometric force-production task. *Experimental Aging Research, 33*(3), 341-353.

van Doorn, R.R., Keuss, P.J. (1992). The role of vision in the temporal and spatial control of handwriting. *Acta Psychologica, 81*(3), 269-86.

Van Emmerik, R.E.A. (1992). Kinematic adaptations to perturbations as a function of practice level in a rhythmic drawing movement. *Journal of Motor Behavior, 24*(1), 117-131.

Van Emmerik, R.E.A., & Newell, K.M. (1989). The relationship between penpoint and joint kinematics in handwriting and drawing. In R. Plamondon, C.Y. Suen, & M.L. Simner (Eds.), *Computer and human recognition of handwriting* (pp. 231-248). Singapore: World Scientific Publishers.

Van Galen, G.P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science, 10*(2-3), 165-191.

Van Galen, G. P., Meulenbroek, R. G. J., & Hylkema, H., (1986). On the simultaneous processing of words, letters and strokes in handwriting: evidence for a mixed linear and parallel model. In H.S.R. Kao, G.P. van Galen, & R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting*. Amsterdam: Elsevier.

Van Galen, G.P., & Schomaker, L. R. B. (1992). Fitts' law as a low-pass filter effect of muscle stiffness, *Human Movement Science, 11*(1-2), 11-21.

Van Galen, G.P., M.M. Smyth, R.G.J. Meulenbroek & H. Hylkema, 1989. The role of short term memory and the motor buffer in handwriting under visual and non-visual guidance. In R. Plamondon, C.Y. Suen & M.L. Simner (Eds.), *Computer recognition and human production of handwriting* (pp. 253-272). Singapore: World Scientific.

Van Galen, G. P., & Teulings H.L. (1983). The independent monitoring of form and scale factors in handwriting, *Acta Psychologica, 54*(1-3), 9-22.

Van Gemmert, A. W. A., & Van Galen, G. P. (1998). Auditory stress effects on preparation and execution of graphical aiming: A test of the neuromotor noise concept. *Acta Psychologica, 98*(1), 81-101.

Van Galen, G.P., & Van Huygevoort, M.A.E. (2000). Error, stress, and the role of neuromotor noise in space oriented behaviour. *Biological Psychology, 51*(2), 151-172.

Van Gemmert, A.W.A. (1997). *The effects of mental load and stress on the dynamics of fine motor tasks.* Doctoral Dissertation, Radboud University, Nijmegen, Holland, 1997. Nijmegen, Holland: Nijmegen Institute for Cognition and Information.

Van Gemmert, A. W. A. (2004, May). Effects of Parkinson's disease on fine motor control. Paper presented at *Motor Control 2004: Trends and Perspectives* (p. 18). Tempe, AZ: Arizona State University.

Van Gemmert, A.W.A., Adler, C.H., & Stelmach, G.E. (2003). Parkinson's disease patients undershoot target size in handwriting and similar tasks. *Journal of Neurology, Neurosurgery and Psychiatry, 74*(11), 1502-1508

Van Gemmert, A.W.A., & Teulings, H.L. (2006). Advances in graphonomics: Studies on fine motor control, its development and disorders. *Human Movement Science, 25*(4-5), 447-453.

Van Gemmert, A.W.A., Teulings, H.L., Contreras-Vidal, J.L., & Stelmach, G.E. (1999). Parkinson's disease and the control of size and speed in handwriting. *Neuropsychologia,* 37, 685–694.

Van Gemmert, A.W.A., Teulings, H.L., & Stelmach, G.E. (1998). The influence of mental and motor load on handwriting movements in Parkinsonian patients. *Acta Psychologica, 100*(1-2), 161-175.

Van Gemmert A.W.A., & Van Galen, G. P. (1996). Dynamic features of mimicking another person's writing and signature. In M. L. Simner, C. G. Leedham, & A. J. W. M. Thomassen (Eds.), *Handwriting and drawing research: Basic and applied issues* (pp. 459-471). Amsterdam: IOS Press.

Van Gemmert A.W.A., & Van Galen, G. P. (1997). Stress, neuromotor noise, and human performance: a theoretical perspective. *Journal of Experimental Psychology: Human Perception and Performance, 23*(5), 1299-1313.

Van Loon, E. M., Masters, R. S. W., Ring, C., & McIntyre, D. B. (2001). Changes in limb stiffness under conditions of mental stress. *Journal of Motor Behavior*, *33*(2), 153-164.

Van Soest, A. J., & van Galen, G.P. (1995). Coordination of multi-joint movements: An introduction to emerging views. *Human Movement Science, 14*(4-5), 391-400.

Verbyla, A. P., Cullis, B. R., Kenward, M. G., & Welham, S. J. (1999). The analysis of designed experiments and longitudinal data by using smoothing splines. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *48*(3), 269 – 311.

Verzani. J. (2005). *Using R for introductory statistics.* Boca Raton: Chapman & Hall/CRC.

Vilageliu, O. S., & Kandel, S. (2009, September). *Early development of writing skills: A longitudinal study with pre-school writers.* Paper presented at the meeting of the 14th Conference of the International Graphonomics Society, Dijon, France.

Vince, M.A. (1948). Corrective movements in a pursuit task. *Journal of Experimental Psychology, 38,* 85-103.

Vincent, N., & Bouletreau, V., & Sabourin, R., & Emptoz, H. (2000). How to use fractal dimensions to qualify writings and writers, *Fractals*, 8(1), pp 85-97.

Viviani, P. (2002).Motor competence in the perception of dynamic events: a tutorial. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: attention and performance* (pp. 406–442), Oxford, UK: Oxford Univ. Press.

Viviani, P., Baud-Bovoy, G., & Redolfi, M. (1997). Perceiving and tracking kinesthetic stimuli: further evidence of motor perceptual interactions. *Journal of Experimental Psychology: Human Perception & Performance; 23*(4), 1232–1252.

Viviani, P., Campadelli, P., & Mounoud, P. (1987). Visuo-manual pursuit tracking of human two-dimensional movements. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 62-78.

Viviani, P. Campadelli, P., & Mounoud, P. (1990). Perceptuomotor compatibility in pursuit tracking of two-dimensional movements. *Journal of Motor Behavior*, *22*(3), 407–443.

Viviani P., & Flash T. (1995). Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1):32-53.

Viviani, P., & Terzuolo, C. (1980). Space-time invariance in learned motor skills In G.Stelmach & J. Requin (Eds.), *Tutorials in Motor Behavior.* Amsterdam: North-Holland.

Viviani, P., & Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience*, 7(2), 431-437.

Volcic, R., van Rheede, J. J., Postma, A., & Kappers, A. M.L. (2008). Differential effects of non-informative vision and visual interference on haptic spatial processing. *Experimental Brain Research, 190*(1), 31-41.

Volman, M.J.M., & Geuze, R. H. (2000). Temporal stability of rhythmic tapping “on” and “off the beat”: A developmental study. *Psychological Research, 63*(1), 62-69.

Von Neumann, J. (1958). *The computer and the brain.* New Haven: Yale University Press.

Vredenbregt, J., & Koster, W. (1971). Analysis and synthesis of writing. *Philips Technical Review, 32*(3-4), 73-78.

Wachter C., Gerisch, H., Navé, P.M.W., Staude, G., & Wolf, W. (2004). Synchronization of periodic movements with external events and interlimb coordination: phase entrainment or phase resetting? *Measurement science review, 4*(2), 20-28.

Wang, Y. (1998). Mixed effects smoothing spline analysis of variance. *Journal of the Royal Statistical Society B, 60,* 159-174.

Wann, J.P., & Nimmo-Smith, I. (1991). Evidence against the relative invariance of timing in handwriting. *Quarterly Journal of Experimental Psychology, 42*(1), 105-119.

Wann, J.P., & Nimmo-Smith, I. (1991). The control of pen pressure in handwriting: A subtle point. *Human Movement Science, 10*(2-3), 223-246.

Watson, R. W., Jones, R.D., & Sharman, N.B. (1997). Two-dimensional tracking tasks for quantification of sensory-motor dysfunction and their application to Parkinson's disease. *Medical and Biological Engineering and Computing*, *35,* 141-145.

Weaver, R. L., Lobkis, O. I, & Yamilov, A. (2007). Entrainment and stimulated emission of auto-oscillators in an acoustic cavity. *Journal of the Acoustical Society of America*, 122, 3409-3418.

Weiner, N., 1948, *Cybernetics: Or the control and communication in the animal and the machine.* Cambridge, MA: MIT Press.

Weisberg, S., & Cook, R. D. (1999). *Applied regression including computing and graphics.* NY: Wiley.

Weisstein, E. W. "Quantile-quantile plots." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/Quantile-quantile plots.html.

Welford, A. T. (1973). Stress and human performance. *Ergonomics, 16,* 567-580.

Weltmann, K.D., Koepke, M. E., & Selcher, C. A. (2000). Spatiotemporal laser perturbation of competing ionization waves in a neon glow discharge. *Physical Review E, 62*(2), 2773 - 2781.

Werner, P., Rosenblum, S., Bar-On, G., Heinik, J., & Korczyn, A. (2006). Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61B,* (4), 228-236.

West, B. T., Welch, K. B., & Gałecki, A. T. (2007). *Linear mixed models: a practical guide using statistical software*. Boca Raton, FL: Chapman & Hall/CRC.

Widla, T. (1990). The influence of spontaneous writing on stability of graphic features. *Forensic Science International, 46*, 63-67.

Wiesendanger, M. (1998). Bernstein's principle of equal simplicity and related concepts. In M.L. Latash (Ed.). *Progress in motor control: Bernstein's traditions in movement studies.* Champagne, IL: Human Kinetics.

Wildgen, W. (2004). *The evolution of human language: scenarios, principles, and cultural dynamics.* Philadelphia: John Benjamins.

Williams, G.P. (1997). *Chaos theory tamed*. Washington, DC: Joseph Henry.

Wimmers, R.H., Beek, P.J., & Van Wieringen, P.C.W. (1992). Phase transitions in rhythmic tracking movements: A case of unilateral coupling. *Human Movement Science, 11*(1-2), 217-226.

Wing, A. M. (1978). Response timing in handwriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 153-172). New York: Academic Press.

Wing, A. (2000). Motor control: Mechanisms of motor equivalence in handwriting. *Current Biology, 10* (6), R245-R248.

Wing, A., & Baddeley, A. D. (1979). A simple measure of handwriting as an index of stress. *Bulletin of the Psychonomic Society, 11*(4), 245-246.

Winstein, C.J., Pohl, P.S., Lewthwaite, R. 1994. Effects of physical guidance and knowledge of results on motor learning: support for the guidance hypothesis. *Research Quarterly for Exercise and Sport,* 65(4), 316-323.

Wishart, L.R., Lee, T.D., Cunningham, S.J., & Murdoch, J.E. (2002) Age-related differences and the role of augmented visual feedback in learning a bimanual coordination pattern, *Acta Psychologica, 110*(2-3), 247–263.

Woch, A., & Plamondon, R. (2004). Using the framework of the kinematic theory for the definition of a movement primitive. *Motor Control*, *8*(4), 547-557.

Wolfram, S. *Mathematica, Version 5.2*. Champaign, IL: Wolfram.

Wooldridge, J. (2008). I*ntroductory econometrics: A modern approach* (4th ed.). Florence, KY: Cengage Learning.

Wood, S.N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, *99*(467), 673-686.

Wood, S.N. (2006). *Generalized additive models: An introduction with R*. Boca Raton, FL: Chapman & Hall/CRC.

Wood, S.N. (May 15, 2009). *Package 'mgcv'.* Downloaded August 16, 2009 from http://cran.r-project.org/web/packages/mgcv/mgcv.pdf.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review* (Monograph Supplement No. 13, Whole).

Wulf G., & Prinz W. (2001). Directing attention to movement effects enhances learning: A review. *Psychonomic Bulletin & Review, 8*(4), 648-660.

Yates, F.E. (1982). The 10th J.A.F. Stevenson Memorial Lecture. Outline of a physical theory of physiological systems. *Canadian Journal of Physiology and Pharmacology, 60,* 217–248*.*

Young, D. E., & Schmidt, R.A., (1992). Augmented kinematic feedback for motor learning. *Journal of Motor Behavior, 24*(3), 261-273.

Yu, H., Russell, D. M., & Sternad, D. (2003). Task–effector asymmetries in a rhythmic continuation task. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 616–630.

Zaki, N. N., & Ibrahim, M. A. (1983). Effect of alcohol and cannabis sativa consumption on handwriting. *Neurobehavioral Toxicology and Teratology, 5*(2), 225-227.

Zelaznik, H. N., Baria, A., Bloom, M., Dolansky, L., Justice, S., Patterson, K., Whetter, E., Spencer, R. M. C., & Ivry, R. B. (2005). Timing variability in circle drawing and tapping: Probing the relationship between event and emergent timing. *Journal of Motor Behavior*, *37*(5)*,* 395-404.

Zelaznik, H. N., Mone, S., McCabe, G. P., & Thaman, C. (1988). Role of temporal and spatial precision in determining the nature of the speed-accuracy trade-off in aimed-hand movements. *Journal of Experimental Psychology: Human Perception and Performance, 14*(2), 221-230.

Zesiger, P., Mounoud, P., & Hauert, C-A. (1993). Effects of lexicality and trigram frequency on handwriting production in children and adults. *Acta Psychologica*, *82*(1-3), 353-365.

Zhai, S., Kong, J., & Ren, X. (2004). Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies, 61,* (6), 823-856.

Zuur, A.F., Ieno, E. N., & Smith, G.M. (2007). *Analyzing ecological data.* NY: Springer.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., & Smith, G.M. (2009). *Mixed effects models and extensions in ecology with R*. NY: Springer.

Zuur, A.F., Ieno, E.N., & Elphick, C.S. (2010, in press). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology & Evolution*. doi: 10.1111/j.2041-210X.2009.00001.x