The Fundamental Issues of Pen-Based Interaction with Tablet Devices

by

Michelle Kathryn Annett

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Computing Science
University of Alberta

Abstract

Although pens and paper are pervasive in the analog world, their digital counterparts, styli and tablets, have yet to achieve the same adoption or frequency of use. Digital styli should provide a natural, intuitive method to take notes, annotate, and sketch, but have yet to reach their full potential. There has been surprisingly little research focused on understanding why inking experiences differ so vastly between analog and digital media and amongst various styli themselves. To enrich our knowledge on the stylus experience, this thesis contributes a foundational understanding of the factors implicated in the varied experiences found within the stylus ecosystem today.

The thesis first reports on an exploratory study utilizing traditional pen and paper and tablets and styli that observed quantitative and behavioural data, in addition to preferential opinions, to understand current inking experiences. The exploration uncovered the significant impact *latency*, *unintended touch,* and *stylus accuracy* have on the user experience, whilst also determining the increasing importance of *stylus* and *device aesthetics* and *stroke beautification*. The observed behavioural adaptations and quantitative measurements dictated the direction of the research presented herein.

A systematic approach was then taken to gather a deeper understanding of device latency and stylus accuracy. A series of experiments garnered insight into latency and accuracy, examining the underlying elements that result in the lackluster experiences found today. The results underscored the importance of visual feedback, user expectations, and perceptual limitations on user performance and satisfaction. The proposed *Latency Perception*

Model has provided a cohesive understanding of touch- and pen-based latency perception, and a solid foundation upon which future explorations of latency can occur.

The thesis also presents an in-depth exploration of unintended touch. The data collection and analysis underscored the importance of stylus information and the use of additional data sources for solving unintended touch. The behavioral observations reemphasized the importance of designing devices and interfaces that support natural, fluid interaction and suggested hardware and software advancements necessary in the future. The commentary on the interaction - rejection dichotomy should be of great value to developers of unintended touch solutions along with designers of next-generation interaction techniques and styli.

The thesis then concludes with a commentary on the areas of the stylus ecosystem that would benefit from increased attention and focus in the years to come and future technological advancements that could present interesting challenges in the future.

Preface

This thesis is an original work by Michelle Annett. The research projects, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, as "Gestures and Interactive Devices", Pro00027121 on January 16, 2012, and as "Evaluation of Human-Computer Interaction Devices and Techniques", Pro00033336 on September 30, 2012. Some of the research conducted for this thesis was performed while I was an intern at Microsoft Research, in Redmond, Washington, USA. Throughout this time, Dr. Walter Bischof was my University of Alberta supervisor, Dr. Anoop Gupta and Dr. Paul Dietz were my Microsoft Research mentors, and Albert Ng worked with Dr. Paul Dietz to design and test the hardware presented in Chapter 3.

Chapter 2 of this thesis has been published as Michelle Annett, Fraser Anderson, Walter F. Bischof, and Anoop Gupta, "The Pen is Mightier: Understanding Stylus Behaviour While Inking on Tablets" in the proceedings of the 2014 Graphics Interface conference. I was responsible for designing the experiment, developing the hardware apparatus, recruiting participants, running the experiment and collecting data, analyzing and interpreting the results and writing the manuscript. Walter and Anoop served as advisors and assisted in the manuscript editing process. Fraser also assisted with various revisions of the manuscript.

Chapter 3 of this thesis has been published as Albert Ng, Michelle Annett, Paul Dietz, Anoop Gupta, and Walter F. Bischof, "In the Blink of an Eye: Investigating Latency Perception during Stylus Interaction" in the proceedings of the 2014 SIGCHI Conference on Human Factors in Computing and in Michelle Annett, Albert Ng, Paul Dietz, Walter F. Bischof, and Anoop Gupta, "How Low Should We Go? Understanding the Perception of Latency While Inking" in the proceedings of the 2014 Graphics Interface conference. For this research project, I was responsible for designing and testing the stylus used in the experiment, designing the experiments, recruiting participants, running the experiments and collecting data, analyzing and interpreting the results, and developing the Latency Perception Model. I also wrote the Graphics Interface manuscript and worked with Albert and Paul to write the SIGCHI manuscript. Albert and Paul designed and implemented the hardware apparatus and Walter assisted in the initial design of the user studies. Walter, Anoop, and Paul were advisors and also assisted in the manuscript editing processes.

Chapter 4 of this thesis has been accepted for publication in ACM Transactions on Computer-Human Interaction as Michelle Annett, Walter Bischof, and Anoop Gupta "Exploring and Understanding Unintended Touch during Direct Pen Interaction". I was responsible for developing the hardware apparatus, designing the experiment, recruiting participants, running the experiments and collecting data, analyzing and interpreting the results, and writing the manuscript. Walter and Anoop were my advisors and assisted in the manuscript editing process.

Chapter 5 of this thesis is in preparation for publication as Michelle Annett and Walter Bischof as "Hands, Hover, and Nibs: Exploring the *Other* Factors Influencing Stylus Accuracy on Tablets". I was responsible for designing the experiments, recruiting participants, running the experiments and collecting data, analyzing and interpreting the results, and writing the manuscript. Walter was my advisor and assisted in the manuscript editing process.

"The art challenges the technology, and the technology inspires the art"

- John Lasseter

Acknowledgements

First off, my advisor Walter F. Bischof deserves my utmost appreciation and gratitude. Walter, you have shared your research experience and psychological knowledge with me on a daily basis, teaching me everything I know about statistics and experimental methodologies (ahem, effect sizes!) Although I still despise your red pen, your attention to detail and inquisitive nature helped shape the writer and researcher that I am today. I know that I would not be anywhere close to where I am if it were not for your high standards and constant pursuit of research excellence. Regardless of how wacky or non-traditional Fraser's and my ideas were (e.g., "We want to build a talking monkey Walter", "Walter, we're going to build a Smell-o-Vision interface"), you never wavered in your support of Fraser and me. You let me spread my wings and go experience industry, first in Toronto, at Autodesk Research, and then again in Redmond, at Microsoft Research. You have always encouraged me to take research risks and immerse myself in projects I enjoyed, no matter how many times I changed my dissertation topic! Although Fraser and I still love our mac and cheese, we will always cherish the culinary adventures you took us on and the magic of a Swiss Raclette! *Eine Minute auf der Lippe, ein Leben lang auf der Hüfte!*

I also would like to extend my utmost thanks to my industry mentors. Dr. George Fitzmaurice and Dr. Tovi Grossman, you both took a chance on an unknown student from the University of Alberta and let her run wild with hot glue and proximity sensors. My time at Autodesk Research opened my eyes to the excitement of industry and underscored the importance of working with great people to me. Thank-you. Dr. Anoop Gutpa and Dr. Paul Dietz, like Tovi and George, you both took a chance on a little Canadian girl and gave her the project and experience of a lifetime. I will never forget my time in Redmond and the opportunities you provided. Paul, thanks for indulging a Disney fan in your many stories and tales about Imagineering and Disney.

There are also many people who I have worked with behind the scenes that have been instrumental throughout my PhD. First off, Dr. Pierre Boulanger, thank you for ensuring that I always had new toys and technology to tinker with and for letting Fraser and me continually rearrange the lab, stock it full of interns, and turn the spring and summer terms into chaos for the sake of "science". I will miss calling the AMMI Lab and CSC 363 my home. I would also like to thank Wei Li, Justin Matejka, Dr. Daniel Wigdor, and John Yee, who were incredibly helpful during my time at Autodesk Research and ensured that Medusa was a success (sorry for disrupting your workspace Justin!). I also owe very big thanks to the members of the Microsoft Research Hardware Lab, including Dr. Tom Blank, Patrick Therien, Chris O'Dowd, and Bruce Cleary. I am very grateful that you welcomed me in the lab and were incredibly patient with my zillions of questions about laser cutting, electronics, 3D printing, and so on.

Much respect and gratitude are also owed to Dr. Abram Hindle, Dr. Brian Maraj, and Dr. Ken Hinckley for their insightful comments on my candidacy and dissertation documents. Your observations were not only invaluable but also ensured that the bigger picture was kept in focus at all times. Thanks also goes to Alberta Innovates, the Association of Computing Machinery, Google and Google Canada, the Government of Alberta's Scholarship Program, the Heart and Stroke Foundation of Canada, the Killam Foundation, the University of Alberta Department of Computing Science, and the University of Alberta Faculty of Graduate Studies for their financial support, many times over, throughout the course of my Doctoral Studies.

Mom and Dad. Can a daughter ever say thank you enough? It was your purchasing of our first computer when I was in Grade 3 that made me realize how important technology would become in our world. Your decision to build a home that was filled with love, dance, books, power tools, hot glue, LEGO, and educational computer games (hello, Learning Company Super Solvers PC games!) instilled within me the importance of asking questions, taking chances, and above all, knowledge. You have always been supportive of my ideas, goals, and dreams, making sure that I had every opportunity to succeed and find happiness. Your love and support of Fraser's and my education has been boundless. I hope that I have made you proud. Thanks also go to Janice and Krista for being supportive and continually finding creative answers to "What is your sister *still* doing in university?"

And lastly, thank you to my best friend, boyfriend, fiancé, and husband, the wonderful Fraser Anderson. Throughout this journey, you have been an incredible and unwavering source of support, wisdom, kindness, and laughter. From reading endless paper drafts and scholarship applications, to being my conference buddy, sounding board, and sharing in my love of Pixar and LEGO, I will never be able to thank you for your unconditional love. You pushed me, every day, to be better. You gave me 1000% of your support when I went off in pursuit of my career. Thank you for allowing my creativity to blossom, my heart to remain happy, and my Vinylmation collection to explode. With you, Duckie, everything really is awesome.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The field of Human-Computer Interaction has long focused on developing techniques and technology to augment our experiences with the physical world. Through scientific experimentation and observation, computer scientists and psychologists have sought to enhance their understanding of users and users' interactions with technology. The knowledge resulting from such experimentation has, in turn, stimulated the design and innovation of novel devices and techniques that support the fluid, natural interaction found in the physical world.

Pen computing is one area where such observations and innovations have been made. In pen computing, users interact with a digital device such as a tablet, using a digital implement that is similar to a traditional pen or pencil. Such implements, or styli, provide a natural method to doodle, diagram, ideate, sketch, and write whilst proving enhanced functionality not possible with traditional pen and paper (e.g., save, search, undo; Read, 2006). As the stylus harnesses the metaphors and motor skills honed from years of experience with pen and paper, it a natural and intuitive implement to use to complete many tasks, including writing, sketching, and annotation (Hinckley & Wigdor, 2002; Hinckley et al., 2010a, b; Sellen & Harper, 2003).

Although pen computing has undergone much innovation, stylus input has yet to be perfected. Thus, the goal of the work comprising this thesis was to identify and explore the fundamental issues influencing stylus interaction with tablets today. To ground the exploration, a specific emphasis was placed on examining and observing natural inking tasks, including writing, sketching, and annotating. The majority of the thesis thus focuses on the identification of those factors that influence digital inking behavior and in-depth explorations of solutions to lessen or overcome such issues.

1.1 Motivation

Over the last decade, tablets have become one of the most popular and fastest growing consumer products. Given their wireless connectivity, portability, and support for touch-based interaction, tablets should be ideal devices for productivity-based activities such as note-taking, sketching, or annotation. Work by Muller, Grove, and Webb (2012), however, demonstrated that tablets are championed for content consumption activities such as gaming, web browsing, social networking, and email, instead of inking-based activities. This pattern of usage was largely attributed to problems inherent in touch interaction, e.g., the fatigue of extended on-screen typing, inadequate touch responsiveness, poor accuracy and precision, and so on.

Many of the problems currently associated with tablet and touch interaction can be solved using implements all users are comfortable with, i.e., pens and pencils. Pens are accurate and precise by their very nature, providing instantaneous feedback in the form of an ink trail. By varying pressure, it is trivial to create lines of varying thickness or intensity. A pen's digital counterpart, the stylus, also shares much of this same functionality, giving it the potential to become an important fixture within the tablet world. A digital stylus is shaped similar to a traditional pen or pencil and has a nib, buttons, and an 'eraser'. On a capacitive device, a stylus emulates the properties of the finger. It could also contain sensors that communicate its position, tilt, or pressure to a digitizer within a tablet. Styli harness one's fine motor control to offer increased precision and accuracy compared to a finger (Cockburn, Ahlstrom, & Gutwin, 2008; Holzinger et al., 2008; Mackenzie, Sellen, and Buxton, 1991; Tu, Ren, & Zhai, 2012; Zambramski, 2011). This precision enables users to perform inking tasks such as note-taking and sketching with ease. If handwriting or diagram recognition is supported, one's writing can be converted to text or one's drawings can be beautified. Unlike the mouse and keyboard, styli afford direct interaction with content, as opposed to physically separated indirect interaction (Brandl et al., 2008). This direct interaction allows users to diagram, work out equations, sketch, create calligraphy, annotate, or sign their name in more natural and fluid manner than with a mouse or keyboard. When coupled with touch, a stylus-enabled tablet should afford efficient bimanual interaction, enabling the dominant hand to perform inking and the non-dominant hand to provide support via manipulation (Guiard, 1987; Hinckley & Wigdor, 2002; Hinckley et al., 2010a, b). Such a combination allows for seamless transfer of natural and intuitive behaviors from the physical to the digital world (Balakrishnan & Hinckley, 1999; Guiard, 1987; Hinckley & Wigdor, 2002; Hinckley et al., 2010a, b) while also supporting the implementation of new tools and interaction techniques (Fitzmaurice et al., 1999, Hinckley et al., 2010a, b; Hinckley et al., 2013; Vogel & Baudisch 2007; Wagner, Huot, & Mackay, 2012; Yoon, Chen, & Gumbretière, 2013; Zhang et al., 2012).

Although styli have had a recent resurgence thanks to Wacom, Samsung, and Microsoft, work on digital styli actually began in the 1950's. In the earliest implementation, MIT developed 'light pens' or 'light guns', which used cathode ray tubes and internal photocells to assist with aircraft stabilization analysis and air defense scenarios

(Gurley & Woodward, 1959). This technology formed the basis for many future systems including SAGE and the TX-0 machines. In 1963, the TX-2 system enabled Sutherland (1963) to develop the first computer drawing application, SketchPad. With SketchPad, a light pen enabled users to interact directly with a screen to sketch, and reposition or modify content. Around the same time, the Rand Tablet, which enabled users to use natural handwriting to specify instructions to a computer (Davis & Ellis, 1964), and the GRAIL system, which supported handwriting and flowchart creation (Ellis, Heafner, & Sibley, 1969), were introduced. Although each of the aforementioned technologies made great strides, their lack of mobility prevented widespread adoption. With the advent of the touchscreen in the 1970's, many portable tablet-like devices such as Kay's Dynabook (1969) were now possible. The commercialization of pen-enabled tablet devices did not begin until the late 80's and early 90's with offerings from GRiD systems (GridPad, n.d.), Go Corporation (EO Personal Communicator, n.d.), Apple (Newton (platform), n.d.) and Palm (Butter & Pogue, 2009). Microsoft even entered the fray with their Windows for Pen Computing software suite (Windows for Pen Computing, n.d.).

Given the recent popularity and long history of styli, one would assume that there are few open questions to answer. However, this is not the case. As previously mentioned, users typically do not use tablets for work or productivity-based tasks. Unfortunately, most research and commercial efforts have not attempted to understand or improve upon the problems that exist. The research problems tackled thus far have largely been dictated by the hardware available, i.e., active styli systems that detect tilt, pressure, and azimuth such as the Wacom Cintiq. With the release of the Wacom Cintiq, much research has focused on pen pressure discrimination and widgets (Hout, Nancel, & Beuadouin-Lafon, 2008; Ramos, Boulos, & Balakrishnan, 2004; Ramos & Balakrishnan, 2003, 2005, 2007; Ren et al., 2008; Suzuki, Misue, & Tanaka, 2010; Xin, Bi, & Ren 2010, 2012; Yin & Ren, 2008), novel user interface widgets incorporating roll, tilt, or azimuth (Bi et al. 2008; Suzuki, Misue, & Tanaka, 2007, 2010; Xin, Bi, & Ren, 2011), bimanual interaction (Balakrishnan & Hinckley, 1999; Hinckley & Wigdor, 2002; Hinckley et al., 2010a, b; Lopes et al. 2001; Matulic & Norrie, 2012), and the integration of haptic elements into a stylus (Giachis, l'Anson & Prytherch, 2001; Kyung & Lee, 2008; Kyung & Park, 2007; Kyung, Lee, & Park, 2007; Lee et al., 2004; Mizobuchi et al., 2005; Wintergerst et al., 2010, 2011).

This thesis places a renewed focus on the stylus, specifically on increasing the understanding and knowledge regarding the natural behaviors, fundamental issues, and basic functionally that are not supported by today's devices. The knowledge and results attained should thus be highly applicable for the devices of today and the near future because an explicit focus has been given to the behaviors and activities that should be supported by default, instead of functionality that could be possible in the future. A brief review of stylus and tablet landscape is detailed next to give context to the results and issues presented herein.

1.2 Background

Some manufacturers have acknowledged the potential usefulness of pens, including styli with many of their newer devices. In the tablet eco-system, some devices such as the Microsoft Surface Pro, Sony Vaio Duo, and Samsung Galaxy Note were specifically designed to support pen-based interaction, while others such as the iPad or Kindle Fire were not. Tablets designed to support pen-computing often have sensing solutions manufactured by one of two companies, Wacom or NTrig. Wacom uses an electromagnetic induction-capable sensor to detect the presence and position of a stylus (Yamanami, Funahashi, & Senda, 1989). This digitizer is located behind the display and emits an electromagnetic signal that is received by a resonant circuit embedded in the stylus. Wacom's approach does not require a battery within the stylus, so there are few limitations on the diameter of the stylus barrel and batteries do not need to be replaced. If both pen and touch are supported, two separate digitizers are required. As the digitizers are placed behind the display, this increases the thickness of a device and adds to the parallax encountered. NTrig, on the other hand, uses a single sensor positioned above the display to support simultaneous pen and touch (Perski & Morag, 2004). With an NTrig setup, capacitive coupling detects touch and a battery-powered stylus transmits electronic signals to the display for the detection of the stylus. Using advanced signal processing, the digitizer reliably distinguishes stylus from touch events. As this requires a self-powered stylus, the weight and thickness of the stylus barrel are greater than those found with systems using Wacom's technology. Both technologies allow for nibs constructed of various materials.

In tablets that include only a capacitive sensing array, such as an iPad or Blackberry Playbook, the inter-sensor spacing is roughly 6 millimeters. This spacing is dense enough to detect the contact area of the average finger but too large for the detection of an averaged size pen tip, e.g., 0.5 - 2 millimeters. Many third parties have created implements such as styli, paintbrushes, etc., with large nibs composed of conductive material such as foams, fabrics, or rubber. These nibs create contact areas that are large enough to be detected on the screen, thus enabling stylus interaction on such tablets (Wu, 2012). As the density of sensor arrays and signal-to-noise ratios improve, the minimum detectable contact area will subsequently decrease, allowing for nibs as thin as sharpened pencils to be used (Sage, 2011).

A number of alternative methods to detect the stylus have also been proposed in the research literature. Systems such as Touch & Write (Liwikcki et al., 2010), FLUX (Leitner et al., 2009), and Aitenbichler et al.'s (2010) work, make use of the existing Anoto pen technology to enable stylus interaction with FTIR-based devices. With the Anoto technology, each pen has an embedded digital camera that detects a specific subsection of a dot pattern printed on the surface of a display. The detected dot pattern helps identify the location of the pen with little need for additional hardware or infrastructure. An approach that looks promising but requires additional hardware is the GaussSense technology (Liang et al., 2012a, b). With GaussSense, a thin array of magnetic sensors is attached to the back of a capacitive touchscreen device and used in combination with a magnetic capacitive stylus. The strong magnetic field emitted by the stylus distinguishes stylus from touch events, determines the pressure exerted by the stylus, senses the tilt and angle of the stylus, and detects the presence of the stylus above the screen. Lopes and colleagues (2001) employed an IR stylus, emitting light at a distinct wavelength, and an IR camera, to track the stylus on the screen. This approach works with diffuse illumination, FTIR, or laser light plane technologies and does not require additional hardware outside the existing touch setup. Although these approaches are innovative, none of them have been shown to exceed the experiences found with traditional commercial digitizers.

1.2.1 Stylus Landscape

Unlike pen and paper, tablets rely on sensors to detect the position of a stylus and calibration procedures to map the sensor space to screen space. The underlying technology in a tablet dictates the requirements and features an accompanying stylus may possess. Although styli come in many varieties, they can be generally classified along two dimensions: the ability of the stylus to be differentiable from touch and the presence or absence of circuitry to support additional functionality. With differentiation, the ability to distinguish between styli and touch events has great influence on user behavior, either supporting natural hand postures and bimanual interaction, or restricting it (Annett et al., 2014a). Stylus circuitry, on the other hand, places many requirements on the aesthetics, cost, and supported functionality. There are generally four categories of styli: *Passive Styli*, *Auxiliary Styli*, *Differentiated Passive Styli*, and *Active Styli* (Figure 1.1). Given the diverse features and properties of these styli, it is not surprising that user experiences are so fragmented and disjoint within the stylus ecosystem (Annett et al., 2014a).

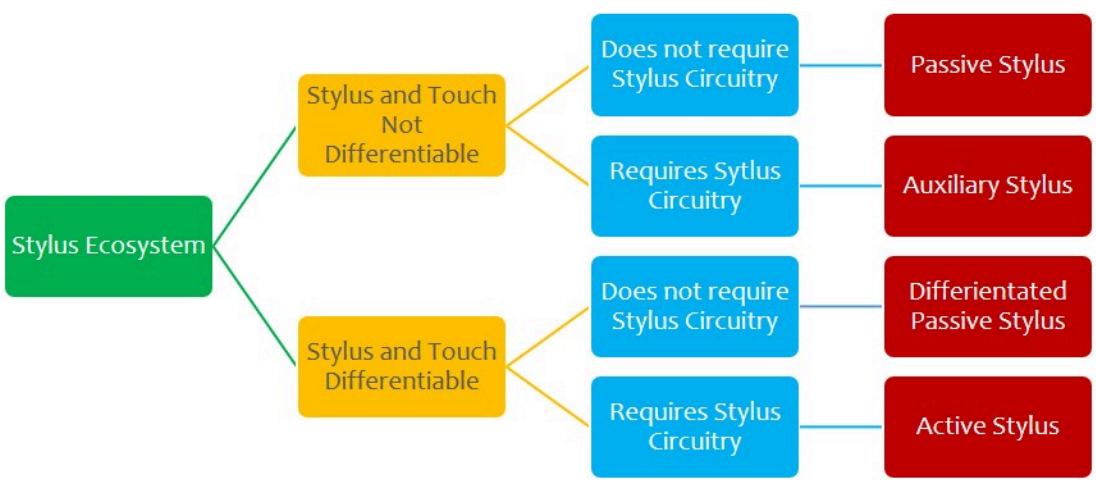


Figure 1.1. The stylus ecosystem today, as categorized by stylus and touch differentiation and the requirement of internal circuitry.

1.2.1.1 Passive Stylus

A passive stylus is one in which the tablet is unable to distinguish the stylus from the finger. Both the barrel and nib of a passive stylus are composed of conductive materials. Such styli are very imprecise, due in large part to the nib dimensions necessary to emulate the imprint of the finger on a capacitive sensing array (Figure 1.2). These nibs are often made of foam, plastic, fiber, or rubber and are not interchangeable or replaceable. Depending on the hardness and deformation of the nib, styli typically have a diameter between 5 and 7 millimeters.

Figure 1.2. Examples of passive styli, from top: Adonit Jot Pro, Wacom Bamboo Solo, New Trent Arcadia Micro-knit, and the Sensu brush.

As these styli do not contain electronic circuitry or sensors, they do not provide information about the angle of the barrel, the pressure exerted while in use, or the nib's location while hovering (Buxton, 1990). Such styli also lack barrel or 'eraser' buttons, thus preventing the erasing functionality common with traditional pencils. Passive styli are popular accessories for capacitive touchscreens such as iPads, the BlackBerry Playbook, Surface RT, or Samsung Galaxy Tab, and are available for a low price (e.g., $5 - $30). Although passive styli are often afterthoughts and introduce many usability problems, they are increasingly being used with current and legacy devices.

1.2.1.2 Active Stylus

With an active stylus, touch input is differentiable from stylus input due to the signal multiplexing performed by a digitizer or through a separate, dedicated sensing array. These styli have very thin nibs, ranging from 1 – 2 millimeters (Figure 1.3), and have optional replacements that modify the friction profile of the nib. This reduces the effects of occlusion caused by the nib and permits the use of natural hand and arm positions. The inclusion of circuitry allows for increased input bandwidth and functionality if barrel buttons or force sensors are included.

Figure 1.3. Examples of active styli, from top: Wacom Cintiq Art stylus, Samsung S Pen, NTrig DuoSense Pen, and the Samsung ATIV Tablet S-Pen.

Such styli detect several levels of pressure, estimate the distance of the pen above the screen, detect the angle of the pen relative to the tablet, and provide feedback regarding the location of the stylus before it touches the screen. Given the electronic circuitry embedded in the barrel, active styli are more expensive, typically $30-$100, but are routinely bundled with products such as the Surface Pro, and Samsung Galaxy Note. While active styli are superior in a number of respects, numerous issues prevent current widespread adoption, such as the increased cost of manufacturing and the limited focus on pen-centric user interaction.

1.2.1.3 Auxiliary Stylus

A new class of styli that has recently become available are auxiliary styli. Such styli are not differentiable from touch, and like the passive styli, they have large conductive nibs made of foam, plastic, or rubber (Figure 1.4) typically 3 to 7 millimeters in diameter.

Figure 1.4. Examples of auxiliary styli, from top: Adonit Jot Touch, Pogo Connect, Pencil and Adobe's Ink (2014).

This class of styli, however, make use of auxiliary input channels, such as Bluetooth, to relay button presses or pressure information from force sensitive resistors embedded within the stylus barrel to a tablet. Due to the wireless connectivity contained within the stylus, they often have larger barrels than passive styli. These styli are almost exclusively available for the iPad and are marketed towards artistic individuals for use in applications such as Sketchbook Pro, Paper, or Photoshop Touch. Due to the Bluetooth and internal sensors, such styli are quite expensive, usually in the range of $80 - $120.

1.2.1.4 Differentiated Passive Stylus

Differentiated passive styli are virtually identical to passive styli, except that they can be distinguished from touch. These styli are compatible with capacitive touchscreens that have optimal signal-to-noise ratios and allow for the discrimination and detection of fine signal fluctuations in the sensor array, enabling robust discrimination of fingers and thin pen nibs. Due to the enhanced signal-to-noise ratios, these styli can have much finer nibs, usually 1 - 3 millimeters in diameter.

This class of styli are largely still in the experimental stage, but a number of prototypes from Synaptics (Coldewey, 2011; Sage, 2011) and Sharp (n.d.; Admin, 2013; Lawler, 2013) appear promising. Given that such advancements are the result of improved capacitive technology, there is little to no additional costs associated with manufacturing them over a traditional passive stylus.

1.3 Thesis Objectives

This thesis presents the results of an exploration into the fundamental issues surrounding pen computing on tablet devices. The main research questions that guided the work included:

- **Chapter 2:** Although pen computing has had a long history, why has it yet to reach mainstream adoption? What are the most prevalent issues that need to be addressed? Are there any behavioral adaptations required today to have a satisfying digital experience?
- **Chapter 3:** As touch and the stylus differ in many ways, what is the minimum latency perceivable when using a stylus? How do these results compare to those found with touch? Can a generalized understanding of latency perception be had?
- **Chapter 4:** What limitations and requirements govern unintended touch? What is the current state of the art and how well do these approaches work while inking? Are there any natural behaviors that, once uncovered, can be harnessed to improve unintended touch? What advancements and directions should be taken to solve unintended touch?
- **Chapter 5:** Aside from occlusion and display parallax, what other factors influence the (perceived) accuracy of a stylus? What effect does visual feedback (Chapter 2), hand postures (Chapter 2), and the design of the nib (Chapter 2) have on accuracy? What is the relative importance of each factor?

1.4 Thesis Organization

The thesis is focused and organized around the research questions detailed in Section 1.3 and largely inspired by the exploratory user study detailed in Chapter 2. The study was designed such that an understanding of the fundamental issues that exist regarding the inking experiences in both the digital and analog worlds could be attained. Details regarding the user study and data analysis are presented, along with the qualitative, quantitative, and behavioral observations that were gathered and synthesized. The synthesis resulted in the identification and exploration of five factors i.e., three primary and two secondary, that most strongly influence the stylus experience today. It is the three primary factors that motivated the experimentation and exploration in Chapters 3, 4, and 5.

Chapter 3 details an exploration into the first of the primary factors influencing the stylus experience today, *device latency*. In this chapter, a comprehensive exploration of pen-based latency perception is presented, detailing three user studies that focused on identifying factors that could influence latency perception while inking. The details of each study, including the technical details of hardware and the psychologically-based experimental methodology, are presented. The results from the three studies, along with relevant prior work, are then synthesized into one of the major contributions of the thesis, the Latency Perception Model. The facets of the model are presented, as are avenues for future explorations of latency and hardware and software advances that can be made.

Chapter 4 presents an in-depth analysis of the second grievance found in Chapter 2, *unintended touch*. After exploring the complexity of unintended touch along the dimensions of timing and data requirements, the nature of input to be rejected, and the cost of recovering from incorrect rejection. The details of a user-based data collection experiment, designed to capture a dataset of natural inking behaviors while participants were writing, tracing, and annotating, is presented. This dataset was then used for an in-depth evaluation of various algorithmic approaches to solve unintended touch. An explanation and analysis of each algorithmic approach is provided first and accompanied by novel behavioral observations that shed light on the benefits and drawbacks of the given approach. The results of a statistical comparative analysis between the various algorithms is also detailed. The chapter concludes with a discussion of factors important when designing unintended touch solutions, including hardware and software improvements, in addition to the importance of harnessing additional data streams.

Chapter 5 focuses on the last primary factor identified in Chapter 2, *stylus accuracy*. The chapter begins by identifying the multitude of sources and factors that influence the (perceived) accuracy of the stylus. Inspired by these factors, the details of two user studies conducted to explore selection and writing accuracy are presented (i.e., one assessing the role of visual feedback and hand postures and the other assessing the design of the stylus nib). The synthesis of the experimental results, along with a relative weighing of the evaluated factors is presented. The implications of the experiments within human-computer interaction are provided at the conclusion of the chapter.

Lastly, Chapter 6 provides a summary of the contributions presented throughout the thesis with regards to the fundamental issues that were identified and explored. It also reviews future research avenues that can be undertaken to advance the field of pen computing and highlights technical innovations that will likely pose continued challenges to pen computing in the future.

Chapter 2

Natural Pen Behavior

Although pen computing has a long history, much of the literature uses active styli to explore bi-manual interaction (Balakrishnan & Hinckley, 1999; Hinckley & Wigdor, 2002; Hinckley et al. 2010a, b; Matulic & Norrie, 2012) or exploit functionality supported by the active stylus such as pressure, tilt, azimuth, and hover (Bi et al. 2008; Grossman, et al., 2006; Ramos & Balakrishnan, 2007; Yin & Ren, 2008)[1]. There is unfortunately little focus or empirical evidence of the problems users face with the multitude of styli available in the marketplace or while *inking*, e.g., drawing and writing. The behavioural adaptations necessary to accomplish routine inking tasks and how to best identify or evaluate inking issues also remains unknown.

To understand the inking experience today, users were observed sketching and note-taking using traditional pen and paper, as well as 'best' and 'worst' case digital devices. While the digital and analog experiences were not believed to be identical, the use of paper provides an excellent gold standard and baseline 'frustration free' inking experience to compare to. The study identified the greatest sources of frustration and poor performance while inking with styli-enabled systems by observing behaviour generated from real world activities, inspecting the content created by participants, and analyzing questionnaires. The resulting factors were then used as motivation for the work presented in later chapters.

2.1 Current Understanding of Pen Behaviour

Throughout the history of pen computing, there has been much work to understand and improve the user experience. Unfortunately, such work has focused on the benefits of styli compared to other input modalities, the tasks styli are best suited for, observed tablet usage and behaviours, and explored pen and paper behaviour.

---

[1]The contents of this chapter has previously been published as: Annett, M., Anderson, F., Bischof, W.F., & Gupta, A. (2014). The Pen is Mightier: Understanding Stylus Behaviour While Inking on Tablets. In the *Proceedings of Graphics Interface,* 193-200.

A number of projects explored usage patterns and behaviour with tablets. Muller, Gove, and Webb (2012) conducted a multi-method exploration of tablet usage and found that activities such as checking email, playing games, and social networking were most common. Wagner and colleagues (2012) explored how users hold a tablet, observing that participants were consistent in the holds used, generally grasping the tablet between their thumb and palm or fingers and palm in consistent locations on or around the tablet. Toy et al. (2012) assessed preferences for web browsing, email, drawing, and gaming on a tablet when the tablet was in the lap, inclined or flat on a desk, finding that most participants preferred to send email and browse the web while the tablet was inclined but preferred the tablet to be flat on the desk when gaming or drawing. These investigations have made important contributions to understanding tablet use and have also underscored the lack of support styli and inking tasks have within the consumer tablet experience. The present work analyses why such a disparity occurs and prioritizes the issues of the stylus experience that are most important and impactful to users.

Many researchers have focused on identifying the activities that the stylus is most beneficial for. Device and task interactions have been largely confirmed, with the stylus identified as optimal for compound tasks, crossing tasks, radial steering, selection, stroke-based gestures, and shape tracing tasks (Cockburn, Ahlstrom, & Gutwin, 2008; Forlines & Balakrishnan, 2008; Holzinger et al. 2008; Mack & Lank, 1989; Tu, Ren, & Zhai, 2012; Zambramski, 2011). Work by Briggs and colleagues (1992) and Ozok et al. (2008) investigated preferences for different tasks and activities supported by a stylus. Activities such as software navigation, pointing, selecting, and sketching were found to be the most preferred, whereas writing incorporating handwriting recognition was the least preferred. Exploratory work by Vogel and Balakrishnan (2010c) observed pointing, selecting, and dragging tasks with a Tablet PC and identified precision, hand occlusion, and the weight and size of the tablet as problematic and frustrating for users. When used with a mobile phone, the stylus also produced more legible notes than a finger (Valderrama Bahmondez et al., 2013). Although styli have yet to be widely adopted, such work illustrates the superiority of styli for many productivity-based tasks and reaffirms the need focus on natural writing and sketching behaviours on tablets, without the use of excessive stroke beautification or handwriting recognition.

Some researchers have also observed pen and paper usage to inform the design and features of digital devices. To inform the interaction techniques used in Manual Deskterity, Hinckley et al. (2010a, b) observed participants' preferred and non-preferred hand usage of a paper notebook, pens, glue, scissors, and paper content while creating a storyboard. Fitzmaurice and colleagues (1999) observed how often participants rotated paper, a tethered tablet, and a 6DOF tablet while completing two handwriting and three drawing tasks, and used the results to inform the design of *rotating user interfaces*. Rosner and colleagues (2008) surveyed existing literature and highlighted the benefits of the physicality of paper, i.e., the folding and dog-earing of pages, the use of dust jackets, and the placement of tabs in notebooks. Lim (2003) explored the differences in thinking processes and cognitive behavior on architects' ability to compose drawings using styli versus pen and paper. It was found that more time was spent inspecting and exploring digital drawings than paper ones, calling for an increased usage of stylus-based systems within the domain of architecture. Oviatt, Arthur, and Cohen (2006) explored how digital devices and pen and paper affected cognitive load and performance, and found that the more interfaces deviate from paper, the greater the cognitive load. Similar to this, the present work observed natural pen and paper behaviour, but in contrast, used their differences to identify the behaviour and adaptations necessary to complete tasks on digital devices. The present work explicitly focuses on understanding the hardware and software features that should be improved to provide a satisfying digital inking experience.

2.2 Experimental Methodology

A user study was conducted to identify the behavioural, performance, and preferential differences between analog and digital media. The exploration was not intended to demonstrate that digital styli provide a better inking experience than pens, as obviously they do not, or prove the superiority of active compared to passive styli, as active is obviously better. Rather, the goal was to understand how various tablet properties affect user behaviour and to design tasks and measures that could be used as a benchmark to determine when the tablet inking experience is 'acceptable' or 'good enough'. Paper was used as a baseline against which to compare the digital tablets because it is the gold standard against which all users (un)consciously compare digital devices and content. The inclusion of a capacitive device which only supported a passive stylus provided verification that the tasks and measures accurately reflected the tablet experience and allowed for an exploration of the full range of behavioural adaptations made with current commercial tablets.

2.2.1 Participants

Thirty participants (10 female) were recruited for the study (*M* = 39 years, *SD* = 10 years, range = 18-60 years). The Edinburgh Handedness Inventory (Oldfield, 1971) classified sixteen right handed (EHI = 73.7) and fourteen left handed (EHI = -57.4) participants. The majority of participants were novice tablet users who had little experience with stylus-enabled devices. Participants were provided with a $10 honorarium for the hour long experiment.

2.2.2 Experimental Apparatus

To better understand the digital versus analog experience, participants used three media: an Apple iPad 2, a Samsung Series 7 Business Slate, and 20 lbs. printer paper (trimmed to 24 x 18 cm). Although capacitive devices are not designed for stylus input, the *iPad* was included as it enabled for an evaluation of the 'worst case' stylus experience and because it has an ever-expanding ecosystem of third party passive and auxiliary styli. The active stylus device, the *Slate*, was specifically designed with stylus support in mind that it is expected to provide a better experience for the user. Non tablet form-factors such as the Wacom Cintiq, Intuos, Cintiq Companion, or Bamboo, were intentionally excluded to allow for an examination of form-factors readily available to everyday users, not niche equipment used by experts and professionals.

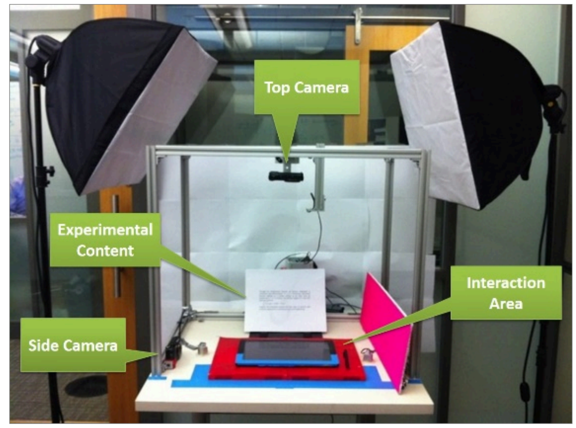


Figure 2.1. The experimental setup with the locations of the cameras, experimental content, and interaction areas highlighted

Three Casio ZR100 cameras were setup to capture participant behaviour (Figure 2.1). One camera was located above the participant and recorded the entire interaction area. As both left and right-handed participants were observed, one of two side cameras captured the vertical movement of the styli and hands.

To enhance ecological validity and repeatability, popular, freely available inking applications were used on the iPad (i.e., *Noteability*) and Samsung Slate (i.e., *PDF Annotator*) instead of custom or professional programs. For inking on the digital devices, participants were provided with a Wacom Bamboo Solo passive stylus with a nib thickness of 7 millimeters and a Samsung ATIV Tablet S-Pen active stylus with a nib thickness of 2 millimeters. Uniball ONYX Fine pens with a nib thickness of 0.7 millimeters were provided for the paper-based conditions. Across all three media, the ink line thickness was approximately 0.7 millimeters and was anti-aliased on the digital devices. Participants were instructed to hold the stylus in their dominant hand and were free to reorient, move, or steady the media as necessary. Stacks of paper were placed under each media to control for the varying thicknesses of the digital devices and paper.

2.2.3 Tasks and Procedure

To elicit natural behaviour and maintain ecological validity, participants completed two activities, *writing* and *sketching*. During the writing task, participants transcribed a paragraph of text on the digital device or sheet of paper containing a mathematical equation (Figure 2.2). An equation was included to capture scenarios where non-traditional, unfamiliar symbols may be used. Participants completed a transcription task rather than generating their own content to ensure behaviour was comparable across media and between participants. It also allowed for divided attention to be imposed, which is common during real-world writing.

With sketching, participants reproduced an organic figure, capturing as many details as possible in five minutes (Figure 2.2). Each figure contained strokes of varying lengths and directions and had explicit shading regions that required both straight and curved lines. In pilot testing, the use of photographs or real objects lead to huge variations in behaviour, with task durations ranging from 30 seconds to 5 minutes. Sketching a unique shape with explicit outlines and shading ensured the task and movements were consistent between participants.

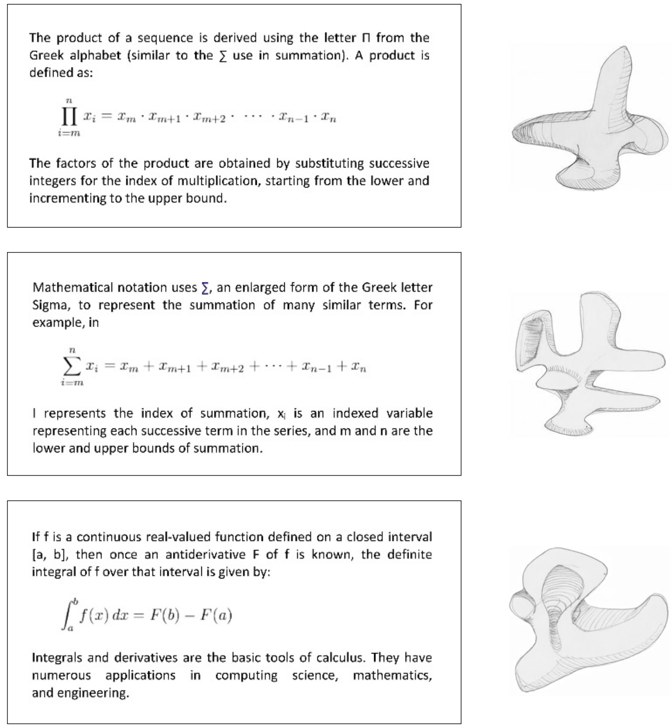


Figure 2.2. The stimuli used in the experiment: The sample paragraphs containing alphanumeric and mathematical content (left) and the organic figures used for the sketching tasks (right).

Each task was completed by every participant on the iPad, Slate, and paper, resulting in six counter-balanced experimental conditions. After each condition, participants answered a questionnaire about their experiences with each medium. Post-experiment follow-up questions were also asked.

2.2.4 Measures and Data Analysis

As it is important to understand behaviour, performance, and preferences, three measures were analyzed: *hand accommodations*, *writing size*, and *user preferences*.

2.2.4.1 Hand Accommodations

Although there are many behaviours that participants exhibit while inking (e.g., rotating and anchoring the tablet, bimanual interaction with dominant and non-dominant hands, and so on) easily identifiable behaviours (i.e., grip and movement style) that had a direct impact on comfort were analyzed. For grip, stylus grips and hand postures previously identified in the literature were consulted (Song et al. 2011; Levy & Reid, 1978). To quantify the hand accommodations, the video data was manually analyzed by one of the authors to find the 'nearest match' for each participant and task. As hand movement styles have yet to be evaluated, the video data was used to identify how the palm and wrist were moving and stabilized while inking. The observed movement patterns were then clustered into groups.

This analysis resulted in 180 unique assignments for both grip and hand movement (30 participants x 3 media x 2 tasks). While a finer-grained analysis reporting on the proportion of time that each participant used each grip or movement style in each task was possible, none of the participants transitioned between grips or movement styles during a task. There were some occasions where a grip was slightly varied (e.g., the fingers were fanned more or less in the crab posture) but the grip remained within the same category during the task. To simplify the presentation of results, grip and hand-movement are provided at a per-task level.

2.2.4.2 Writing Size

The most appropriate measure found to quantify the effects of each device's characteristics on performance during a pilot study was writing size. To compute writing size, the average height from the baseline to x-height (Font Shop, 2010) of each line was computed and averaged at three different points along each line (not including the larger sigma or pi characters). The writing size of the line containing the equation (i.e., *equation line*) was computed separately from the other lines, (i.e., *text lines*) resulting in two measures of performance. Writing size measures are presented in points to be consistent with common typographic conventions; twelve points was equivalent to 4.2 millimeters, as measured on-screen.

2.2.4.3 User Preferences While Inking

As perception and opinions towards devices are important, ratings regarding the appropriateness of each medium were collected for writing and sketching. Participants were asked to indicate if they felt that "the {Paper, Slate, iPad} was a good medium to complete the {sketching, writing} task with" on a 7-point Likert scale. A freeform comment section asking "Why did you provide this rating?" and "What did you like and dislike about this medium?" gathered deeper insights into the ratings provided by participants.

2.3 Results

For each measure, the observed behaviours are detailed and followed by the statistical analysis and synthesis of the results. Where appropriate, Bonferroni corrections were applied to the post-hoc, pairwise comparisons. Note that the analyses were designed to provide insights into the features affecting tablet use. By determining which behaviours were significantly different, and to what degree, the relative importance of various tablet properties could be understood. As the study was exploratory in nature, external rather than internal validity was favored. Definitive claims regarding universal properties of active or passive styli is not claimed, as there are many differences between the implements' underlying digitizing technology.

2.3.1 Hand Accommodations

Across all experimental conditions and participants, each observed grip and movement pattern was classified and a multinomial logistic regression analysis determined if any grips or movement patterns were unique to specific media, tasks, or the handedness of participants.

2.3.1.1 Grip

Prior work has identified a variety of grips common during content creation tasks (e.g., writing, drawing, painting, tool use, and so on). Levy and Reid (1978) identified one grip, the *natural* grip, and described an 'inversion' behaviour that some subjects, especially those that were left handed, performed. Work by Selin (2003), identified a number of pen grips that adults and children employed while writing with pen and paper, including the *radial cross palmar grip*, the *palmar supinate grasp*, the *digital pronate grasp*, the *brush grasp*, *grasp with extended finger, thumb tuck*, *index grip*, and *thumb wrap*. She also found the *tripod*, *relaxed tripod*, and adapted *tripod* grips, which were variants of the *natural* grip found prior by Levy and Reid. While observing four participants use Wacom devices, Song et al. (2011), observed five pen grips, i.e., *tripod*, *relaxed tripod*, *sketch*, *tuck*, and *wrap*. These five grips were found to be largely task dependent (e.g., the wrap grip used only for painting, the sketch grip to shade with a pencil lead, etc.) and were similar to the grips previously identified by Levy and Reid and Selin.

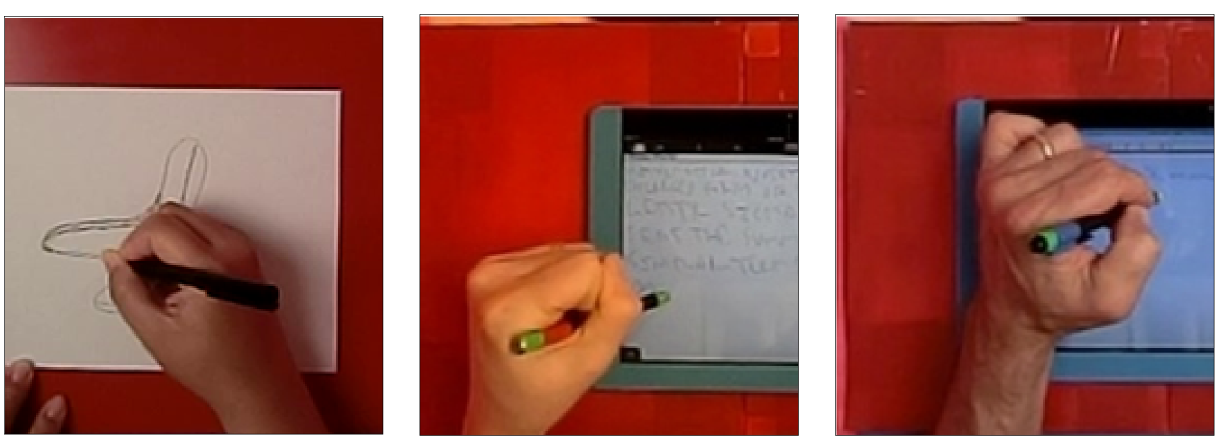


Figure 2.3. Examples of the grips exhibited by participants. From left: the natural grip, the knuckle grip (knuckles curling in to grip the pen), and the crab grip (knuckles fanned out to support the hand).

Of the multitude of grips identified in the literature, only the *natural* grip, was observed during the experiment (Levy and Reid, 1978; Selin, 2003; Song et al., 2011). Two other grips, the *knuckle* and *crab* grips (depicted in Figure 2.3, usage in Figure 2.4) were novel. Across all tasks and media, the most frequent grip, *natural*, was exhibited 61% of the time. The second most common grip, used 23% of the time, was the *knuckle* grip. With this grip, the knuckles were aligned parallel to the top of the tablet and clenched around the barrel of the pen. The least common grip was the *crab* grip (12%), in which participants fanned the fingers not gripping the stylus, similar to a crab's legs. The remaining 4% of the grips were assigned to a fourth, *other* category.

A multinomial logistic regression, using *natural grip* as the reference category, evaluated the influence of Medium, Handedness, and Task on grip. None of the factorial interactions influenced the makeup of the model so they were removed (i.e., Device x Handedness p = .781, Task x Handedness p = .941, Device x Task p = .920, and Device x Task x Handedness p = .105). In the resulting regression, Device ($p < 0.05$), Task ($p < .001$) and Handedness ($p < .001$) influenced the grip used. While writing, participants were less likely to use a crab grip than while sketching ($p < .01$). Left-handed participants were more likely to exhibit a crab grip ($p < .05$), knuckle grip ($p < .001$), and other behaviours ($p < .05$) than right-handers. Additionally, participants were more likely to use an unclassified, 'other' behaviour with the passive stylus than with the active stylus or on paper ($p < .001$).

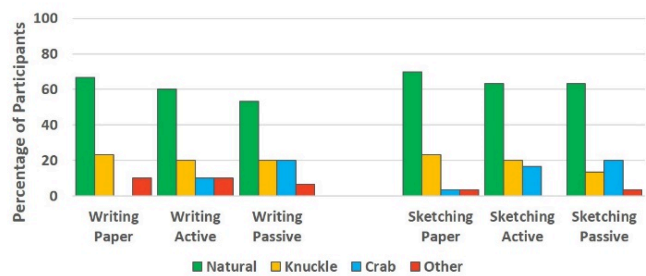


Figure 2.4. The grips displayed by participants while writing and sketching, presented by task and medium. Note the prevalence of the 'crab' grip while using digital devices.

Across all tasks and media, the natural grip was overwhelmingly the most popular, albeit slightly less prevalent with the digital devices. Interestingly, the novel crab grip used almost exclusively with digital devices was widely reported by participants as a method to overcome unintended touch. As passive stylus systems are more prone to stray marks (due to the lack of stylus sensing), the prevalence of crab grips with the passive system illustrates how widespread the problem of unintended touch is. Unlike active stylus systems, passive stylus systems cannot predict where the stylus will be, so unintended touch becomes much more difficult and almost *requires* the use of hand accommodations when unavailable.

The grip analysis also demonstrated that left-handed participants were more likely to use the knuckle grip than right-handers. Although none of the participants exhibited the inverted, or hooked, grip often associated with left-handed writers (Guiard & Millerat, 1984; Levy & Reid, 1978), those that exhibited the knuckle grip also rotated the digital media 20 - 40 degrees. The identification of the knuckle grip and the accompanying medium rotation is important for designers of orientation-aware widgets and developers of unintended touch solutions to note, as the palm would likely produce a very different pattern on digitizer than with the natural grip.

2.3.1.2 Hand Movement Style

Participants exhibited one of three categories of hand movement patterns: *floating*, *planting*, or *dragging* (as depicted in Figure 2.5). The most prominent behaviour, *floating*, was exhibited by 51% of participants. While floating, participants held their wrist, palm, and/or fingers aloft, above the writing surface. The second most popular pattern of movement was *planting* (39%), whereby participants planted their hand on the surface and wrote or sketched until the current word or stroke was complete. Participants then picked up their hand, moved it to a more convenient location, and replanted it on the screen. The least frequent behaviour was *dragging* (10%), where the hand was placed on the media and drug across the surface until it reached the end of the line or the stroke being made. At this point, it was picked it up and moved it to the next location.

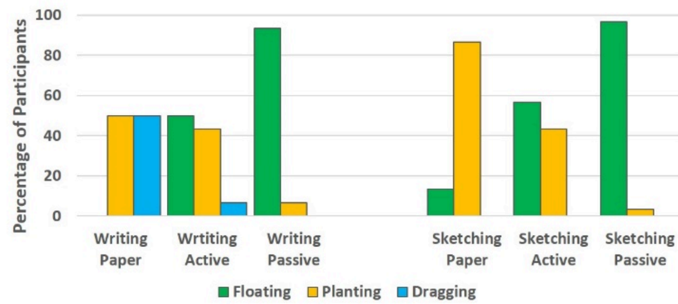


Figure 2.5. The patterns of hand movements used by participants while writing and sketching. Note the lack of floating with paper and the lack of dragging with digital devices.

Another multinomial logistic regression was performed using the *planting* behaviour as the reference category to determine the role of medium, task, and handedness on hand movement. None of the factorial interactions were found to influence the makeup of the model so they were removed (i.e., Device x Handedness $p = .212$, Task x Handedness $p = .713$, Device x Task $p = .687$, and Device x Task x Handedness $p = .999$). The resulting regression revealed that Device ($p < .001$), Task ($p < .001$), and Handedness ($p < .05$) all influenced the movement of the hand. Participants were more likely to use a dragging movement on paper than the digital devices ($p < .01$), and more likely to use the floating behaviour with the passive than active stylus ($p < .001$) and active stylus than paper ($p < .001$). Left-handed participants were also less likely to use a dragging behaviour than right-handed participants were ($p < .05$).

Hand dragging was used almost exclusively on paper, with only two participants dragging their palm on the active system while writing. On paper, participants reported that they were able to slide their hand along the page much easier because the friction between their hand and the surface was suitable. On the digital devices, however, the level of friction was too high, leaving many participants unable to slide their hand naturally.

Although participants were encouraged to interact normally and were told that they could rest their palm, almost all participants modified their behaviour when using the passive stylus. The difference in the floating movements with passive versus active styli suggests that the active system's method to identify and reject unintentional touch events was slightly better than on the passive system. When touch events were improperly handled, many more extraneous touch points were created than participants were comfortable with, so they used a different movement style, i.e., the floating behaviour. The frequency of planting on paper and the active tablet indicates that the digital devices supported the transfer of normal writing behaviours and hand postures. It is possible that those who lifted the palm when using the active stylus were pre-conditioned to lift their palm by prior experiences with passive styli or other touchscreen devices.

2.3.2 Writing Size

To evaluate the character size used while writing (Figure 2.6), a mixed-design ANOVA was conducted, with Device (levels: paper, passive, active) as the within-subjects factor and Handedness as the between-subjects factor (levels: left, right). Handedness was not found to be significant (Text lines: $F_{1,27} = 0.1$, $p = .788$; Equation Line: $F_{1,27} = 0.2$, $p = .652$), so the handedness factor was collapsed and another ANOVA was performed without this factor. This second ANOVA determined that Device influenced the writing size (Text lines: $F_{2,50} = 10.0$, $p < .001$; Equation Line: $F_{1.6,39.5} = 12.8$, $p < .001$). Post-hoc pairwise comparisons determined that participants wrote smallest on paper, slightly larger with the active stylus, and largest with the passive stylus (Table 2.1). When writing the equations, participants wrote substantially larger than while writing the text lines. As the equation lines contained characters that were presented and are often written larger, it is somewhat expected that this behaviour was transferred to the digital devices.

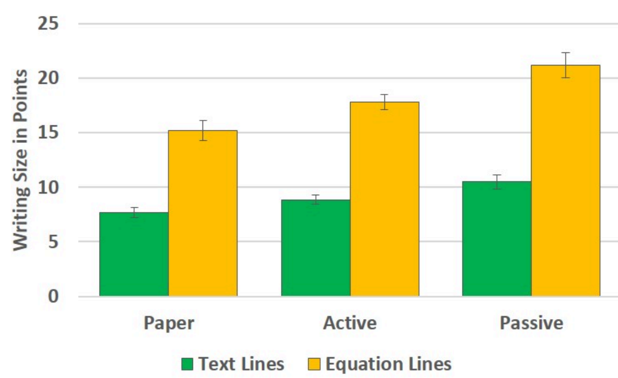


Figure 2.6. The writing size used on each device when writing the text lines and equation lines. Note the increase in size from paper, to active and passive. The error bars represent the standard error of the mean.

Table 2.1. The results of the writing size pairwise comparisons (* denotes significance).

Post-hoc Comparisons	*p* for Text Lines	*p* for Equation Line
Passive vs. Paper	.001 *	.01 *
Paper vs. Active	.01 *	.25
Passive vs. Active	.05 *	.05 *

The increased writing size, from paper to active to passive, is indicative of the accuracy differences that exist between the media. Many participants believed that the passive system was "incapable of detecting any strokes smaller than a ¼ inch so (they) had to write and draw much larger than normal". With the active device, the precision of the nib and feedback provided about the presence and location of the nib before it touched the screen enabled participants to write at sizes very close to that of paper. This aided in the perceived accuracy of the Slate (e.g., "you can see the pen tip before you touch the tip on the surface") and improved the stylus experience, "I could actually put content where I wanted."

2.3.3 User Preferences While Inking

No significant differences were found between the handedness groups with respect to the Likert-scale ratings (Figure 2.7), so handedness was collapsed and a Friedman's ANOVA was performed. Participant's opinions towards each device were found to be significantly different ($p < .001$). Wilcoxon-signed rank post-hoc tests revealed significant differences for each media (Table 2.2), with paper being the most preferred, followed by the active stylus and Samsung Slate and then the passive stylus and iPad.

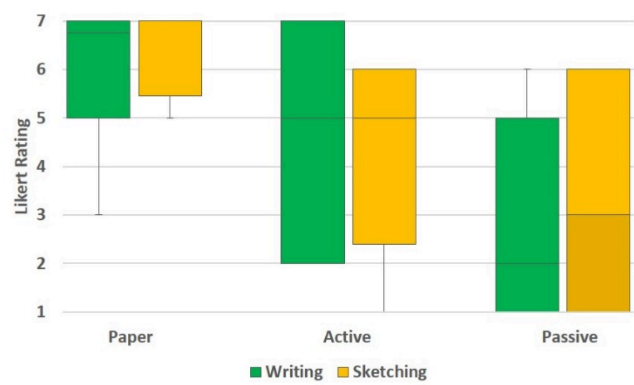


Figure 2.7. Participant median responses to "I feel that the {Paper, Slate, iPad} was a good medium to complete the {sketching, writing} task with". Note the decline in ratings from paper, to active and passive.

As expected, paper was preferred by all participants. As it has zero latency, a natural feel and texture, provides direct contact with no parallax, is lightweight and easy to manipulate, palms glide easily across its surface, ink flows easily across the page, and the nib provides audio feedback as it scratches the paper's surface, it is the gold standard. In contrast, the passive device received very poor ratings (i.e., a median response of 'Mostly Disagree'). Passive stylus systems are not designed for productivity-based tasks; this was reflected in participants' ratings. As active stylus systems are optimized for inking, it is reaffirming that the active system was rated higher than the passive system (i.e., the median response was 'Slightly Agree').

Table 2.2. The results of the Wilcoxon post-hoc analysis for the questionnaire data collapsing across Handedness (* denotes significance).

Post-hoc Comparisons	*p* <
Sketching passive vs. Sketching paper	.01 *
Sketching passive vs. Sketching active	.05 *
Sketching paper vs. Sketching active	.01 *
Writing passive vs. Writing active	.01 *
Writing passive vs. Writing active	.01 *
Writing paper vs. Writing active	.01 *

The passive system was also rated slightly higher for sketching than writing. Although participants disliked the passive system for sketching, it is interesting that they felt it was slightly more appropriate than for writing. As sketching is inherently a messy task, the less accurate movements may have masked the passive system's other deficiencies. It is also possible that the perception many had towards the iPad as a 'finger painting' device also had an influence, along with the visual affordances of the passive stylus, which looks similar to a marker (e.g., "the marker-like, thicker pen gave me a much better drawing experience").

The subjective responses echo what was seen with the objective measures: paper provides the best inking experience, followed by the active stylus and device, and lastly the passive stylus and device. While this is expected, this supports the validity of the tests and measures used.

2.4 Discussion

The experiment drew attention to, and uncovered, many elements that influence behaviour, performance, and preferences for digital versus analog media while inking today. From the observations and participant comments, many features emerged, with five presenting substantial, pressing issues (Figure 2.8). Participants were most vocal about three *primary features*, i.e., stylus accuracy, device latency, and unintended touch, and *two secondary features*, i.e., stylus and device aesthetics and stroke beautification. The prioritization of these features was based on the number of comments each received in addition to the behavioural and performance impact found with the quantitative results. The identification these primary and secondary features, as well as differences between the tablet devices, warrants much further work and has inspired the focus and research that constitutes the remainder of this thesis.

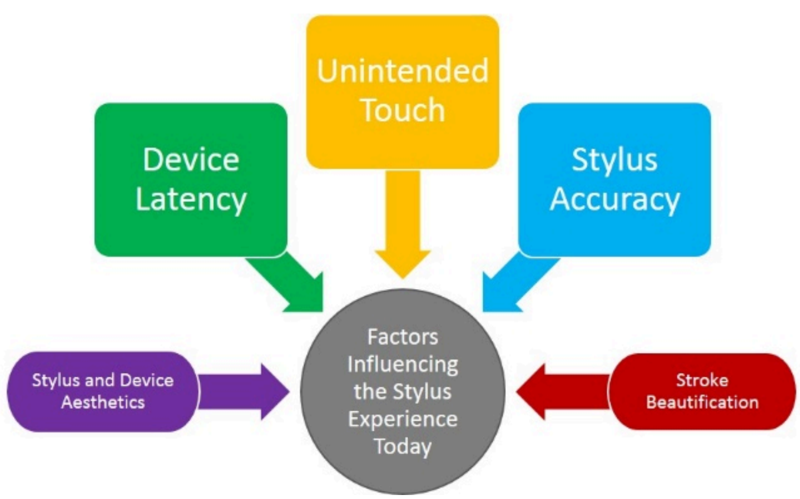


Figure 2.8. The five factors identified as being most impactful to the stylus experience today. The major factors (i.e., device latency, unintended touch, and stylus accuracy) are explored in more detail throughout this thesis.

2.4.1 Primary Features

Device latency, stylus accuracy, and unintended touch affected participant behavior the most, resulting in the greatest effect on participant behaviour and were the most prominent features identified by participants as problematic.

2.4.1.1 Device Latency

One hundred milliseconds has long been regarded as the minimum latency necessary for satisfying direct-interaction (Miller, 1968). Although the iPad and Slate had end-to-end latencies below this threshold (80 and 65 milliseconds respectively, computed using the method in (Ng et al., 2012)), many participants commented on the sluggish nature of the tablets, noting that "the digital ink did not flow naturally from the stylus". The delayed ink and inaccuracy caused participants to write slower and larger to "see what (they) had already written so that (they) could better join the parts of each letter together instead of having to guess where they would be because of the delay". Although the active system was only 15 milliseconds faster than the passive system, some participants appreciated the difference, stating that "the (active) pen tracked as fast as I was able to write which was great" and "the Slate was much faster than the iPad, but it was of course still slower than paper". Such comments and the increased writing size corroborate with recent work on touch-based latency perception wherein increased latency decreased performance (Jota et al., 2013). While the experiments demonstrated that latencies as low as 25 milliseconds had an impact on performance, manufacturers are a long way from achieving such latencies with commercial products.

2.4.1.2 Unintended Touch

The digital devices prevented participants from interacting naturally because participants altered their behaviour to avoid making unintended, accidental markings. Such markings were due the tablets being unable to distinguish between the intended and unintended touch events, i.e., deliberate touch actions versus those caused by resting one's palm or grazing the fingers over the surface. With the passive system, participants were "forced to write in an uncomfortable position to avoid the 'palm touch' screen" and "could not rest (their) palm on the display without disrupting it – highly unusable". With the active system, participants were "more willing to interact because (they) could rest (their) palm on the surface with no problems" and "the Slate didn't have the palm 'touchy' problems that the iPad did". As they were only inking for 5 minutes, participants did not experience much fatigue. Inking for longer periods would have likely exacerbated fatigue and the issues associated to unintended touch.

Some manufacturers have acknowledged the importance of unintended touch. While recent devices tote 'palm block' or 'palm rejection' technology, in practice, such implementations are far from robust, detecting many spurious touch points. Unintended touch will continue to be a problem whenever both pen and touch are supported, regardless of if they are used synchronously (e.g., bimanual interaction) or asynchronously (e.g., interleaved interaction). Future work should focus on understanding the behaviors and stylus features that can be harnessed to improve rejection and prevent the adaptions observed in the present study.

2.4.1.3 Stylus Accuracy

A recurring theme that emerged was frustration due to inaccuracy of the stylus (i.e., the stylus did not deposit ink where the user expected it would). Many participants were vocal about inaccuracy, as it forced them to alter their writing size and speed. Participants had more difficulty forming and terminating letters with the passive system than the active system or paper (Figure 2.9). The feedback available to the user when the active stylus was in the hover state (Buxton, 1990), mitigated the effects of inaccuracy and provided a much more enjoyable experience, "I loved the pen, and I have never used such an accurate pen before". Inaccuracy also manifested itself while sketching, where many participants made larger, straighter, seemingly haphazard strokes with the passive and active systems compared to paper (Figure 2.9).

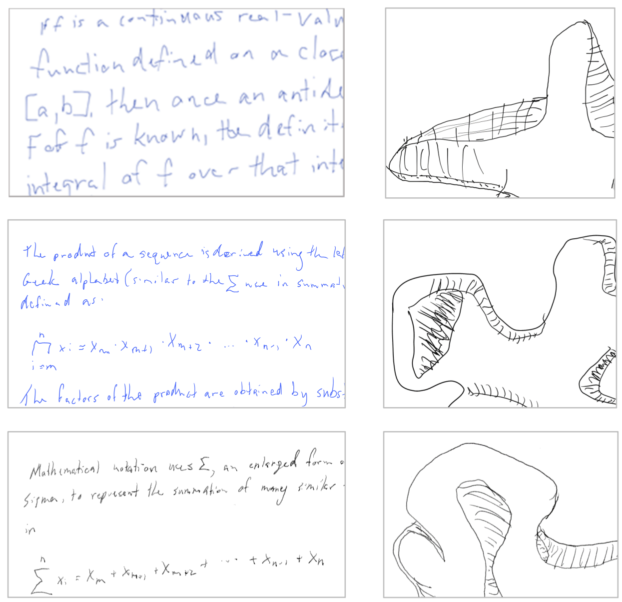


Figure 2.9. Inking content from the same participant created using the passive stylus (top), active stylus (middle), and pen and paper (bottom). The images were cropped to show detail, and the paper image was scanned, resulting in the perceived loss of quality. Note the inaccuracy of the lines on both digital devices, as well as the limited use of curved lines with the passive device.

The composition of the digital styli also affected the perceived accuracy and precision of strokes, with the passive styli being perceived as less accurate, "I couldn't see where I was writing because of the (passive) squishy pen so I had to write bigger" and "I couldn't tell where the lines would start or end". With the active system, however, many participants believed that "the pen tip felt almost like a real pen" and many felt that it mimicked a traditional pen quite well. Although parallax (Lee et al., 2012; Ramos et al., 2007; Vogel & Balakrishnan, 2010a) and occlusion (Annett et al., 2014a; Badam et al., 2014; Ramos et al., 2007; Vogel & Balakrishnan, 2010a) influence device accuracy, the physical design of the stylus may be implicated as well.

As inaccuracy forced participants to write larger, in the real world, it would subsequently result in less content fitting on the screen. It is important to consider the design implications of this. As screen real estate is already constrained by menus and UI elements, the area available for content creation is at a premium. It is thus important to consider the benefits that intuitive navigation methods robust to users resting their hand or fingers on the screen, and intelligent widgets or canvases capable of reflowing or reformatting content (Bargeron & Moscovich, 2003; Breuel et al., 2002; Brush et al., 2001; Yoon et al., 2003), could have if they become integrated within stylus-supported applications.

Although accuracy has long been a complaint of tablet users, it is still a problem. Based on comments and the quantitative results, accuracy appears to be influenced by many factors including the ability to receive feedback in the hover state, the size of the nib, cursor calibration, the texture of the screen, the material composition of the stylus, and the responsiveness of the device. Although third-party auxiliary styli such as Ink, Pencil, and Adonit's Jot Pro have begun to use alternative input channels to provide some of these features to passive systems, there remains much work for designers of both passive and active systems. Further experimentation is still needed to tease apart these factors, in addition to many other factors including sensor linearity (Ward & Philips, 1987), transducer parallax (Ward & Philips, 1987), and so on, to better understand and minimize inaccuracy.

2.4.2 Secondary Features

Stylus and device aesthetics and stroke beautification were also noted as being important to the stylus experience, albeit to a lesser extent than the primary features were.

2.4.2.1 Stylus and Device Aesthetics

Interestingly, many participants had marked opinions on the texture of the digital devices and stylus aesthetics. Influenced by years of writing on paper with pens, users are accustomed to specific tactile sensations and feedback while holding a stylus. While inking with the digital devices, there was a discrepancy between the friction of the hand and surface. Many participants felt that "there was not enough friction between the pen and screen to feel natural" and "my hand jerked across the screen as I moved it". This mismatch was also reflected in the number of participants who floated their palms above the surface of the tablets. The importance of surface texture to participants counters current thinking about tablet surfaces, i.e., that they should be made of glass because it is glossy, slick, and visually appealing. If stylus-based devices are to be taken seriously as pen and paper replacements, the texture of a device's surface and the materials the stylus nib is composed of should be optimized to evoke familiar feedback patterns and encourage natural movements instead of hindering them.

The aspect ratio of the active device was also problematic for some. A 16:9 ratio is well suited for watching movies, but insufficient for writing notes and sketching. Some participants mentioned the lines they wrote on the Slate were "going on forever and ever" and that they had to "squish (their) sketches to fit on the Slate but not the iPad". Unlike Vogel and Balakrishnan's research (2010a), none of the participants mentioned the increased thickness or weight of the digital devices compared to paper. This is likely because our participants did not have to support the devices themselves while performing the tasks.

A few participants also noted that end-user customization and choice is important in stylus designs. Current styli come in muted colors (i.e., black or grey) and the choice of nibs is limited (except with Wacom active styli). Compared to traditional pens that come in a myriad of shapes, sizes, weights, and ink types, (e.g., gel, ballpoint, felt-tipped, fountain), digital styli feel impersonal. Having the opportunity to customize the stylus and appearance of ink would invoke a stronger connection to one's work, which is a natural benefit of writing with a pen. Additionally, a variety of after-market gloves or surface coverings could be designed, allowing users to choose the texture they prefer, similar to the assortment of nibs available for Wacom styli.

2.4.2.2 Stroke Beautification

Many participants commented on the appearance of their strokes. Applications today often modify ink thickness, opacity, and path smoothing using input parameters such as pressure, velocity, and time to imitate real ink dynamics. With the passive system, participants identified that the stylus was not pressure sensitive and were unhappy that this feature was not supported, especially while sketching, e.g., "the lack of pressure sensitivity is annoying" and "without pressure sensitivity, the strokes looked awful". Although the active system made use of a pressure-sensitive stylus and anti-aliased strokes, none of the participants believed it was pressure sensitive. Although many have developed stroke beautification techniques (Fekete et al., 1995; Lu et al., 2012; Zitnick, 2013), the current beautification methods employed for inking did not meet participant's expectations.

These observations highlight the importance of beatified strokes and the value of appropriate rendering techniques. Even if an application "fakes it", users want the illusion of pressure sensitivity, "that Paper app has pressure and I know that it's fake but I still enjoy it". As the cost of styli become cheaper and it becomes easier to integrate auxiliary communication channels into passive styli, designers should re-evaluate the role of pressure, tilt, azimuth, and barrel roll in ink rendering outside the context of pressure-based widgets, novel interaction techniques, or the levels of tilt, azimuth, or pressure discernible that have been the focus as of late (Bi et al., 2008; Vogel & Balakrishnan, 2010; Xin, Bi, & Ren, 2010, 2011, 2012). Such improvements will uphold beliefs that tablets can provide experiences similar, if not more appropriate and engaging, to pen and paper for productivity-based tasks (Oviatt, Arthur, & Cohen, 2006).

2.5 Summary

Although pen computing has had a long history, little information has been available regarding the varying inking experiences in the analog and digital worlds. The study presented in this chapter provided evidence of the adaptations and behaviours that occur while performing inking tasks on tablets today. By comparing these behaviours to those observed with traditional pen and paper, grips and patterns of hand movement unique to digital devices and left-handed users were identified. These behaviours, as well as device characteristics, resulted in larger characters when writing, inaccurate strokes, and user frustration.

The experiment identified the major features influencing the inking experience today. Stylus accuracy, device latency, device and stylus aesthetics, digital ink rendering, and the ability to distinguish between intended and unintended touch are of the utmost importance and warrant much more attention. Although the devices used in the present study were not the most recent available on the market, they still represent the state of the art in terms of tablet experience. Device latency, surface texture, and unintended touch have not seen significant advancements in recent years. The tablet and stylus have great potential to become 'go-to' devices for inking and productivity-based activities, but many improvements are needed before they become commonplace. In lieu of these issues, the remainder of this thesis seeks to explore and understand the role the primary features have on the stylus experience today: Chapter 3 focuses on device latency, Chapter 4 focuses on unintended touch, and Chapter 5 examines stylus accuracy.

Chapter 3

Device Latency

As identified in Chapter 2, one of the foremost grievances regarding stylus-enabled devices involves the responsiveness, or latency, encountered while inking (Annett et al. 2014a)[1]. Regardless of the technology used to detect the stylus, the end-to-end latency on most commercial tablets is between 65 and 120 milliseconds (Ng et al., 2012). This is significantly different from pen and paper, where latency is essentially zero because ink instantly flows from the nib onto the page. Latencies are much higher on digital devices because input information and feedback travels through a complex pipeline composed of input sampling from the digitizer, the filtering and reporting of samples to the operating system, the processing samples by the operating system, the reporting and processing of events by an application, and the updating of the display (Hinckley & Wigdor, 2002). Each step in this pipeline adds to the delay users perceive.

As it is unlikely that stylus-enabled devices will ever achieve true zero latency, an achievable goal is to minimize latency such that users have the illusion, or perceive, that a device is more responsive than it truly is (Seow, 2008). To users, the perceived latency would thus seem identical to true zero latency. If users perceive two different latencies as being equal, (e.g., 1 millisecond and 40 milliseconds), it may not be necessary to allocate resources to achieve the latency levels recommended previously by Ng et al. (2012). Although there is an ecosystem-wide push to make systems faster, it may thus be beneficial to reallocate CPU or GPU cycles to improve stroke rendering, integrate pressure or tilt information from the stylus, or to improve the rejection of unintended touch (Chapter 2). Delaying ink by 10 or 15 milliseconds, for example, may be acceptable if the user experience can be improved in other ways.

---

[1]The contents of this chapter has previously been published within: i) Ng, A., Annett, M., Dietz, P., Gupta, A., Bischof, W.F. (2014). In the Blink of an Eye: Investigating Latency Perception during Stylus Interaction. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1103-1112. and ii) Annett, M., Ng, A., Dietz, P., Gupta, A., Bischof, W.F. (2014). How Low Should We Go? Understanding the Perception of Latency While Inking. In the *Proceedings of Graphics Interface*, 167-174.

To understand latency within the context of real-world activities and work towards an understanding of the factors that influence latency perception, three experiments were conducted. The aim of this work was not to determine how latency influences performance or develop methods to reduce hardware or software delays, but rather to understand human perception during real world activities. Using a prototype high performance system, users' perception of latency while performing a number of stylus-based tasks was investigated. The results from the experiments, in addition to those from prior work, provided insight into how latency is perceived and helped to form the *Latency Perception Model*, which provides a blueprint for future explorations into latency perception.

3.1 Current Understanding of Latency

Although delays and lag have been a complaint of technology for years, very little work has assessed its influence on the user. Latency research outside virtual reality and video games has focused on direct-touch input. Work by Anderson, Doherty and Ganapthy (2011) used two touchscreen devices to determine the effects of touch latencies between of 80 to 780 milliseconds during everyday tasks such as reading, photo browsing, and web browsing. They found that delays of less than 580 milliseconds were acceptable to users. In work by Ng and colleagues (2012), a low latency touch device assessed the latency perception while performing a dragging task. The device consisted of an opaque touch sensor interfaced with a TI Discovery 4100 DLP development kit. The kit's FPGA performed all of the processing and directly controlled a DMD at very high frame rates, allowing a baseline latency of one millisecond. It was found that users were unable to perceive latencies below six milliseconds (range 2 - 11 milliseconds). In follow up work by Jota et al. (2013), the influence of latency, target location, and target width was examined. Performance was found to decrease as target width decreased, latency increased, and target distance increased. The just-noticeable difference while perceiving the 'land-on' portion of the dragging event was 64 milliseconds. These varying results suggest that task may have a great effect on the perception of latency. The device used in this prior work was unfortunately not suitable for a stylus because the sensor resolution was designed for larger input modalities such as fingers and no disambiguation between the hand and a stylus was possible. A new low latency prototype specifically targeted towards stylus input was thus developed.

Although much further work can be done within the touch domain, the present experiments focus on exploring latency perception while using a stylus. Just as with direct-touch input, few have assessed stylus latency, largely due to technical limitations within digitizers and tablets. In early work with a light pen system, Miller (1968) estimated that 100 milliseconds of delay would be acceptable to users while drawing slow deliberate strokes. Unfortunately, few details are available regarding the derivation of this estimate. In recent work, Henzen and colleagues (2004, 2005) developed a low latency electronic ink display for drawing and animation applications.

The display had a minimum latency of 40 milliseconds and exhibited zero parallax. The setup was a prototype and unfortunately did not undergo experimental evaluation with participants.

There has been a variety of work focused on the detection and understanding of latency. Researchers within computational music have strived to determine the ideal latency for musical composition. Early work by Freed, Chaudhary, and Davila (1997) and Wright and Brandt (2001) suggested that music controllers should have less than 10 milliseconds latency, as it is at this point that piano players notice delays. Many others have recommended much higher latencies, depending on the type of instrument, genre of music, and the presence of tactile feedback. Maki-Patola and Hamalainen (2005) determined that delays between 2 and 30 milliseconds were sufficient while playing a Theremin without tactile feedback. Adelstein et al. (2003) found that delays of 24 milliseconds were tolerable while tapping a brick with a hammer. Dahl and Bresin (2001) recommended latencies of 55 milliseconds for percussion instruments. While playing collaborative music via network connection, delays of 100 milliseconds were acceptable while playing piano, but only 20 milliseconds while playing an accordion (Sawchuk et al. 2003). Using this work as a guide, Montag and colleagues (2011) built a low-cost multi-touch tabletop capable of providing low latency audio and haptic feedback to users for musical applications. The system improved audio-haptic synchrony, achieving a minimum latency of 30 milliseconds. The fragmented results and recommendations from the computer musical literature demonstrate that many factors, i.e., feedback modality (e.g., tactile, audio, or visual feedback), task, and input, influence the perception of latency during musical composition (Maki-Patola, 2004).

3.2 High Performance Stylus System

To determine the minimum perceivable latency, the prototype High Performance Stylus System (HPSS) system from (Ng et al., 2014) was used (Figure 3.1). The HPSS employed two Texas Instrument Discovery 4100 high-speed projector kits, a first-surface mirror for rear-projection onto a diffuse surface, and a fiber-optic stylus. The Discovery kits were able to achieve high frame rates using Digital Micromirror Devices (DMD). DMD's contain arrays of micromirrors that modulate light very quickly, allowing binary frames to be projected at a rate in the tens of kHz. The first projector kit rear-projected a series of grey-coded patterns that utilized Lee et al.'s (2004, 2005) structured light technique. The IR grey-coded patterns were projected at 17,000 frames per second at a resolution of 1920 x 1080. The patterns uniquely encoded every pixel in the image area. To provide visual feedback, a second Discovery 4100 kit refreshed at 23,000 binary frames per second and had a pixel resolution of 1920 x 1080.

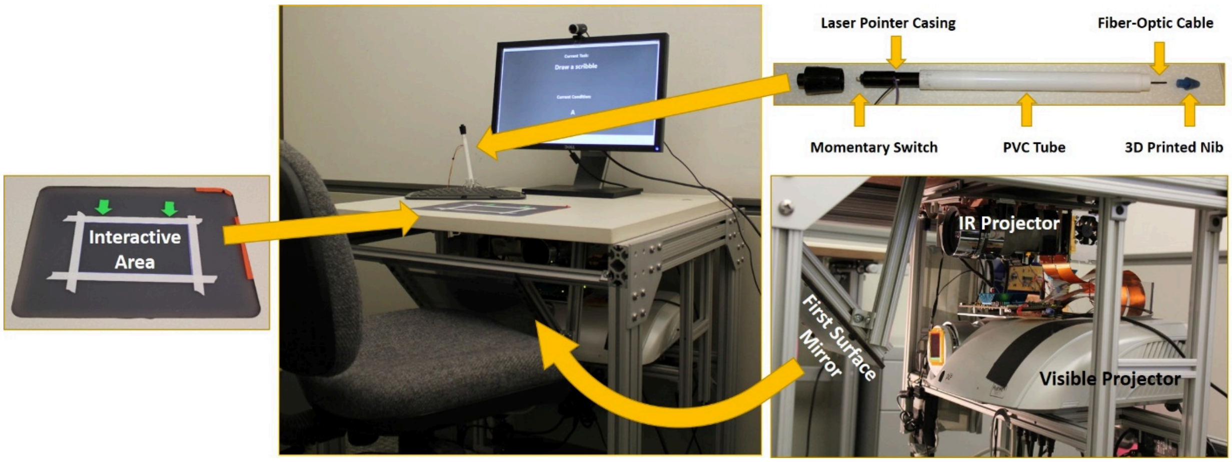


Figure 3.1. The prototype High Performance Stylus System, composed of two high-speed projectors, a first surface mirror for rear projection, and a fiber-optic based stylus. Further details about the hardware can be found in (Ng et al., 2014).

The stylus used a one-millimeter fiber optic cable to detect the grey-coded IR patterns. As users vary greatly in the way they hold a stylus, it was important that the stylus could be robustly actuated at a variety of angles. The stylus thus underwent many iterations of design, finally converging on a design that used a one-millimeter flexible fiber optic cable. The cable was bundled within a stylus constructed from PVC tubing, a cap, and a hollowed out laser pointer (Figure 3.1). The fiber optic cable was placed inside the laser pointer casing, and the casing was nested inside the PVC tubing. Attached to the end of the laser pointer casing was a momentary switch. Whenever the stylus was pressed or removed from the screen, the laser pointer (de)pressed the switch against the cap, switching the ink on or off. This simple construction allowed participants to write with the stylus at any angle comfortable to them.

As the fiber optic cable was very thin, the nib of the stylus was 3D printed from UV cured ABS plastic and a one-millimeter hole was drilled in the tip. The diameter of the nib's tip was 1.2 millimeters. The resulting stylus was 187 millimeters long, had a 13-millimeter barrel diameter, and weighed 19 grams, which is close in weight and size to a typical Wacom stylus.

Although the High Performance Stylus System is capable of running independently, a HP Z400 Workstation was connected to the system via a serial connection to manipulate the latency values. A custom C# and WPF program automatically determined and sent appropriate latency values to the system, gathered latency judgments from participants, and recorded the minimum perceived latency for each task. A 21" Dell monitor provided participants feedback about the current task and condition and prompted them for their latency judgments. To advance to the next condition, and indicate their latency decision, participants pressed the A, B, and space bar keys on a Microsoft Arc keyboard.

3.3 Just Noticeable Difference Methodology

As the goal of the present work was to determine the minimum latency perceivable, it was thus appropriate to use a just-noticeable difference (JND) methodology. With JND methodologies, participants are presented with two stimulus levels and are forced to make a judgment regarding which alternative was brighter, quieter, faster, and so on. After repeated presentations of various stimuli, one can derive the minimum threshold, or *just-noticeable difference* (JND), that is perceivable for a given stimulus. The JND paradigm converges on a threshold that is the result of participants being unable to distinguish the minimum baseline from all latencies below the converged threshold. Because the task is held constant across trials, it is assumed that participants would be unable to distinguish between any latencies lower than the threshold.

During the experiments, two latencies were presented on each trial, the *baseline,* which was held constant and acted as a reference for the participant, and the *test,* or *probe,* that was modified on each trial. Although many methods can be used to determine the test value, it is important to choose method that reflect the needs of the experiment. Prior work used staircase methods that have been around since the inception of psychophysics (Ng et al., 2012; Jota et al., 2013). Given that the experiments required repeated motor movements, a highly efficient method that mitigated fatigue and increased engagement was needed. The more modern Parameter Estimation by Sequential Testing (PEST) adaptive technique (Taylor & Creelman, 1967) met these requirements. This newer methodology produces little variance in the resulting thresholds compared to legacy methods, allows the experiment to be completed faster (30-80 trials), and reduces participant fatigue and boredom. The duration of an experiment using PEST is approximately 10 minutes.

With PEST, the Wald (1947) sequential likelihood-ratio test uses the prior history of a participant's responses at a given stimulus level determined the test latency and the amount that the stimulus should increase or decrease by (step size). Once the step size reaches a minimum, i.e., 1 millisecond, the experiment concluded (aka McMillian and Creelman's Minimal Overshoot and Undershoot Sequential Estimation technique (1991)). This ensured that participants experienced the smallest possible difference between latencies. An upper bound was placed on the probe latency (i.e., 105 milliseconds) to prevent participants from experiencing levels of latency that exceeded those found today that are already identified as bothersome. If the probe ever reached this level, the experiment concluded.

To further increase the efficiency of PEST, an initial step size of 10 milliseconds and expected probability of 75% were used, i.e., participants would have to correctly identify the baseline latency on 75% of trials before advancing to a lower level of latency. In an ideal scenario, the baseline latency would be zero milliseconds, but as with all prior work, current technology is unable to achieve such latencies. As such, the minimum latency (i.e., baseline) possible with the HPSS was 1 millisecond. As the onscreen digital ink required filtering and smoothing via a moving average window, the minimum latency (i.e., baseline) while inking was seven milliseconds. Motivated by prior work (Ng et al., 2012, Jota et al., 2013), the initial testing latency was set to 65 milliseconds to prevent participants from completing too many trials that would likely be too easy. Across all trials, the presentation order of the baseline and testing latencies were randomized.

3.4 Experiment 1: Latency Perception While Dragging and Scribbling

As little is known about the perception of latency while inking or using a stylus, three psychophysics experiments were conducted. Each was designed to determine the lowest latency detectable using a task hypothesized to increase latency perception, i.e., dragging and scribbling. Such tasks ensured ecological validity and comparability to prior work.

3.4.1 Participants

Sixteen naïve individuals (4 female) participated in the study (*M* = 33 years, *SD* = 9 years, range = 24-52 years). All participants were right handed and had normal or corrected-to-normal vision. Participants had varying levels of exposure to tablets and styli, from complete novices to others who worked with stylus-enabled devices each day. In a pre-experiment questionnaire, thirty-eight percent of participants were familiar with latency, through playing video or mobile phone games or through interacting with virtual environments Each participant was provided a $10 honorarium for the 30-minute experiment.

3.4.2 Tasks

Three tasks were used to determine the lowest latency participants could perceive: *large box dragging, small box dragging*, and *scribbling*.

In the *large box dragging* task, participants placed their stylus on the left hand side of the screen underneath the left arrow, dragged the stylus laterally to the right hand side underneath the right arrow, and then dragged the stylus back underneath the left arrow (Figure 3.2). While doing this, a 20 millimeter x 20 millimeter white box was continually centered at the location of the nib. To maintain consistency with prior work, the same box dimensions were used (Ng et al., 2012).

As the size of the nib was much smaller than the finger, a variation of the box dragging task, *small box dragging*, was also conducted. By closely matching the size of the box to the dimensions of the nib, it should have been

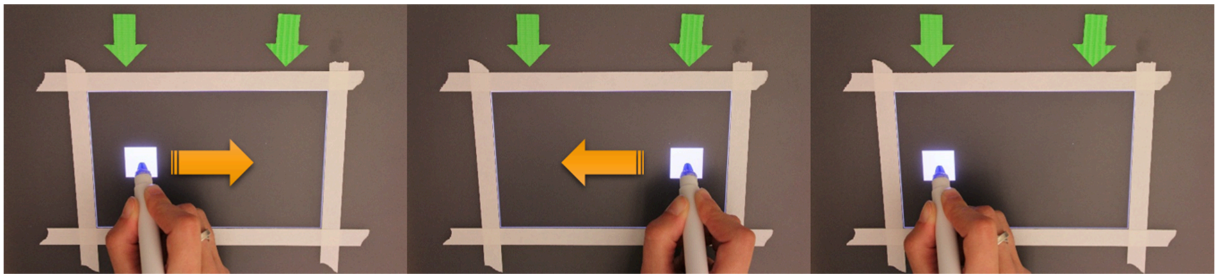


Figure 3.2. The box dragging task used in Experiment 1. Participants were asked to move the box from under the left arrow to the right arrow and then back under the left arrow.

much easier for participants to notice if the nib fell outside the box or was not in the center, and hence enhance their ability to perceive latency. To compute the appropriate box dimensions for this condition, the finger-to-box ratio used by Ng et al. was computed using the average width of the index finger at the proximal interphalangeal joint, i.e., 16 millimeters (Dreyfuss & Tilley, 2002), and the large box dimensions of 20 x 20 millimeters. Using this ratio and the size of the nib, a 6.25 x 6.25 millimeter box was appropriate for this condition.

As we were also interested in latency perception while inking, in the *scribbling* task, participants drew a curvilinear line, or scribble, starting in the upper left corner and moving down towards the lower right corner (Figure 3.3). To encourage the same motor movements across all participants, each participant was instructed not to touch the edge of the interactive area with their scribble nor to draw a shallow, wavy line. Feedback was provided in the form of a 1-pixel thick on-screen ink trace. Similar to the dragging tasks, participants were verbally encouraged to maintain the same speed throughout all trials. A scribbling motion was used in lieu of writing a word, performing a free-form sketching task, or tracing a simple shape, as we wanted to encourage the same type of quick, ballistic motions that were elicited during the box dragging tasks.

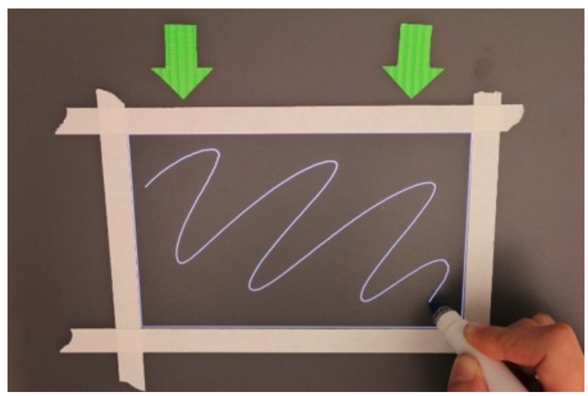


Figure 3.3. The scribbling task used in Experiment 1, wherein participants drew an oscillating line from the upper left corner of the screen down towards the lower right corner.

As the absolute minimum latency users could perceive was of interest, participants were not explicitly required to complete each trial at a specific speed. The experimenter did monitor each participant via web camera and provided verbal feedback to participants if they appeared to be moving at unnatural speeds. As per the requirements of any JND paradigm, participants were required to perform the same task on each trial. To test different shapes or stroke directions, participants would have had to complete another JND experiment.

3.4.3 Procedure

At the start of each experiment, participants sat in an adjustable drafting chair in front of the HPSS. The concept of latency was explained to each participant to ensure that they understood the purpose and goals of the experiment. Participants were then informed that they would be performing a number of inking tasks and that the minimum latency they could perceive would be measured. On each trial, participants were informed that two different latencies, A and B, would be presented. Participants were asked to complete the task twice, first at latency A, then at latency B. To switch from A to B, participants pressed the space bar. After each trial, participants used the A and B keyboard keys to indicate "which condition exhibited the least delay".

Although explicitly priming participants for latency could influence their behaviour, it was imperative to do so. The use of ambiguous questions probing which condition was 'most preferred' or 'better' would have left too much room for interpretation, indirectly encouraging some participants to focus on other factors or visual cues while making their decisions. It should be noted that any values determined are likely to be higher in a real-world scenario, where latency detection is not paramount in a user's mind.

3.4.4 Results

As the baseline latencies for the dragging and scribbling tasks were different (i.e., one millisecond versus seven milliseconds), the two tasks could not be statistically compared. As such, two separate analyses were conducted.

A Wilcoxon signed-rank test compared the two box sizes and revealed that participants were able to discriminate between lower latencies when the smaller box (Mdn = 2 milliseconds) compared to the larger box (Mdn = 6 milliseconds) appeared around the nib, $z = -2.8$, $p < 0.01$, $d = 1.03$ (Figure 3.4). Users were thus able to perceive minute differences caused by latency, similar to results found prior by Ng et al. (2012). The different threshold values suggest that the dimensions of visual feedback in reference to the physical input likely played a role in participant's perception of latency. Dimensions that complement the relation between the physical reference and digital feedback likely provide better visual cues and assist in latency perception.

The results from the scribbling task demonstrated that participants were able to discriminate between the 7-millisecond baseline latency and a median of 40 milliseconds (Figure 3.4). Although not directly comparable to the dragging tasks, such results suggest that task demands may play a role in the perception of latency. The higher perceived latency found while scribbling compared to dragging may be due to the different visual feedback available, the strategies used to determine latency, or the cognitive load encountered while scribbling versus dragging the box.

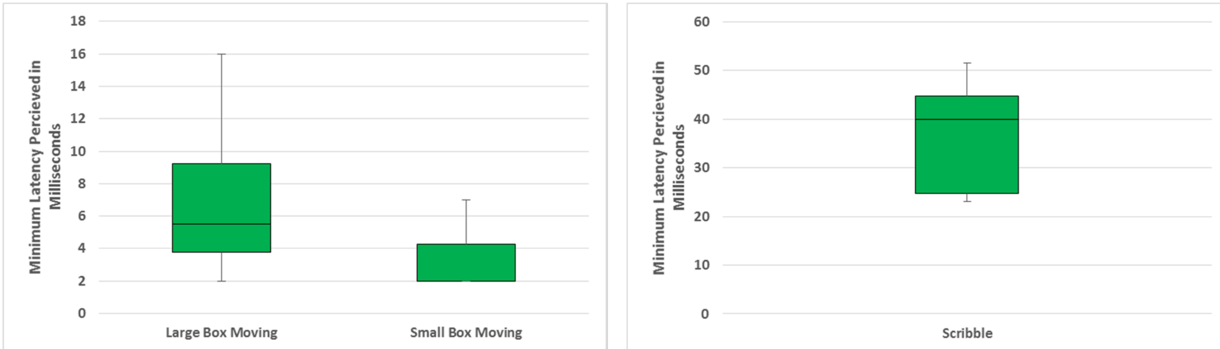


Figure 3.4. The median latency perceived by participants while dragging the large and small boxes (left) and performing the scribbling task (right). The box boundaries represent the 5th and 95th percentiles.

3.4.5 Discussion

The results found with the dragging and scribbling tasks provide valuable information about the basic visual process and the ability of participants to detect and notice latency. Given the latencies possible on devices today, i.e., 65 - 120 milliseconds, these results additionally provide support for decreasing latency along all levels of the pipeline.

Users appear to detect latency by attending to the visual disparity between a visual reference, which may move, and corresponding feedback available from the display. The perceptibility of this disparity appears to depend on a number of factors. The size and type of feedback presented likely influenced participant's comparisons. With the scribbling task, a thin persistent line was visible, whereas with the dragging tasks, a large box whose position constantly updated was visible. As there was minimal occlusion from the stylus nib, it is likely that participants had an easier time viewing feedback as the box(es) activated a larger area within the foveal or parafoveal region. In the original latency experiments by Ng et al. (2012), the finger occluded a larger area of the visual feedback than the stylus nib, so this may explain why participants found it easier to detect latency, especially when the dimensions of the physical input were appropriately scaled to the feedback. When considering Jota et al.'s (2013) tapping task, it is likely that the haptic feedback generated by touching the screen was integrated with the visual feedback from the display to make latency judgments. The combination and synchronization of feedback from other modalities likely decreases the perception of latency, hence the higher thresholds found by Jota et al. Further work is thus needed to truly understand the role of feedback in the perception of latency.

As participants were free to use whatever method necessary to make their latency judgments, it is also possible that the location of focus and subsequent visual cues available influenced perception. When asked how comparisons were made, participants reported different focus loci for the dragging versus scribbling tasks. While dragging, participants reported fixating on the box's edges to determine if the nib was inside, outside, or in the center of the box, which is likely similar to what occurred in the prior touch work by Ng et al (2012). As there was high contrast on the screen, and little occlusion from the nib, such a determination was easier with the stylus than with touch. The smaller the box, the easier it was to make this judgment. While scribbling, the location of focus was not always on the nib. Some participants reported focusing on the ink, others on the pen or hand and some continually switched between the ink and the pen. Such strategies changed the location of the nib and ink on the retina as well as the distance between them. While varied, such attentional patterns made different visual cues available to participants and fall in line with the eye tracking literature (Alamargot et al., 2006; Coen-Calgi et al., 2009; Gowen & Miall, 2006; Miall, Imamizu, and Miyauchi, 2000).

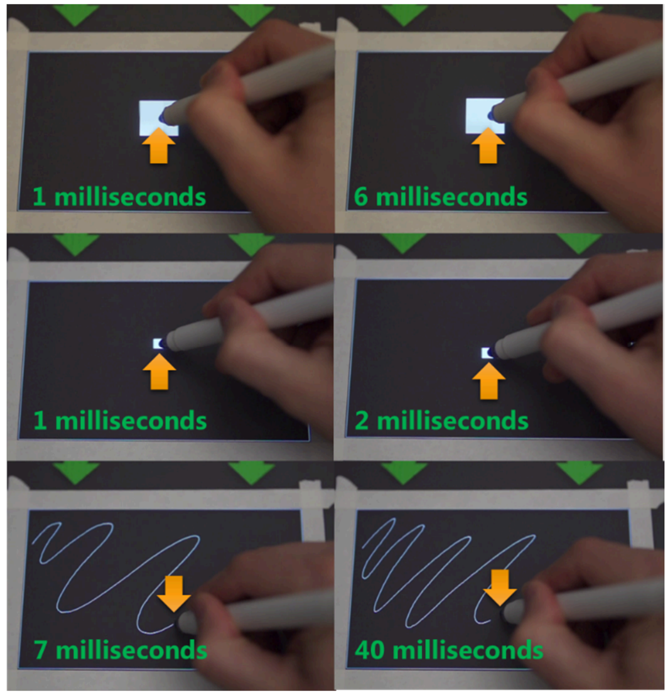


Figure 3.5. A demonstration of the latencies perceivable by participants, where the yellow arrow indicating the current nib location. The images were recorded using a high speed camera at 480 fps. From top: The nib is to the right of the box during the large box dragging task, slightly off center during the small box dragging task, and there was a visible gap between the nib and ink while scribbling.

Cognitive load and the attention required to complete a specific task may also influence latency perception. The tapping and dragging tasks used in this and the prior touch work had low levels of cognitive difficulty and did not divert much attention from the latency judgments being made. Repeatedly drawing the same pattern on the screen and trying to maintain its shape and length (scribbling), on the other hand, is more cognitively challenging and attention diverting because participants had to focus on maintaining the same shape and stroke speed on each trial. Although the scribbling task is a rudimentary form of inking, it is likely that tasks with even more cognitive and attentional demands such as taking notes, solving mathematical equations, or sketching a portrait, would divert more cognitive resources away from latency perception, decreasing perception even more. This, of course, requires further investigation. Based on the initial results, it may be reasonable to target latencies closer to 40 milliseconds for stylus-based experiences, instead of the 10 milliseconds recommended prior for touch-based tasks.

While the results provide many insights into the perception of latency, they are not without their limitations. Approximately 63% of participants could discern between one and two milliseconds of delay. When asked if they were guessing or confident in their comparisons, most indicated that they could still discriminate between the two conditions, but it was difficult to do so. There was a very small difference in the location of the box for one versus two milliseconds, compared to one versus six milliseconds (Figure 3.5). As the human eye can discern motion or deviations in spatial patterns separated by one arc minute (1/60th of a degree), if the HPSS could display latencies of finer fidelity, e.g., 1.5, 1.6, etc. or the baseline latency could be lowered below one millisecond, it is likely that most participants would be able to discriminate between latencies even lower than one millisecond while performing dragging tasks.

The results demonstrate the minimum latency users can detect, but under conditions where users were explicitly primed to look for latency and do not have attentional competition. In situations where explicit cueing is not available, the minimum noticeable latency will likely be higher. Although the results suggest that developers and device manufacturers have much further work before latency is no longer a complaint from users, they should be considered carefully.

3.5 Experiment 2: Perceived Latency While Inking

As Experiment 1 suggested, a number of factors influence the perception of latency. To further understand latency perception during scenarios that require increased cognitive and attentional demands (compared to the tasks used in the prior experiment), a second experiment determined the minimum latency perceivable while participants preformed real world inking tasks including drawing simple lines, writing, and sketching. As such tasks require a variety of different movements, levels of attention, and produce a variety of visual cues (e.g., simple lines, curves, corners), they were very representative of the feedback and demands users would likely face in the real world.

3.5.1 Participants

Twelve individuals (3 female) with normal or corrected-to-normal vision were recruited to participate in the study (*M* = 34 years, *SD* = 7 years, range 23-44 years). All participants were right handed and were naive to the purpose and goals of the experiment during recruitment to remove any bias or experience with latency from pen-enabled systems, touch enabled systems, or video games. A range of participants were recruited, some used tablets and styli daily and whereas others had limited prior exposure. In a pre-experiment questionnaire, twenty-five percent of participants were familiar with latency, through playing video or mobile phone games or through interacting with virtual environments. Each participant was provided a $10 honorarium at the conclusion of the 30 experiment. None of the participants from Experiment 1 participated in Experiment 2.

3.5.2 Tasks

Three inking tasks were chosen based on their similarity to real world activities: line drawing, writing, and drawing.

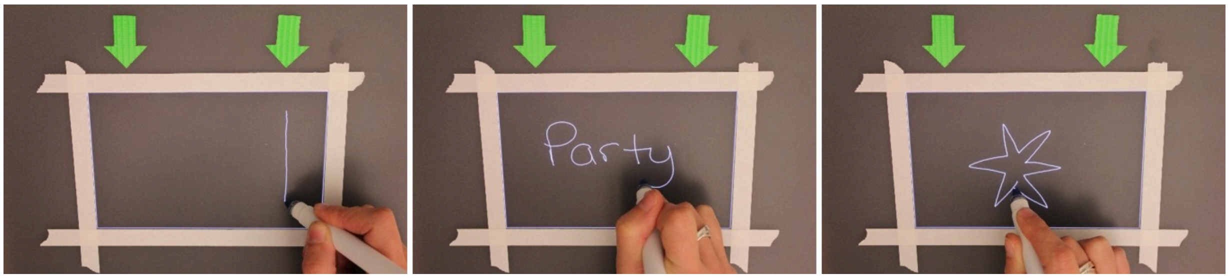


Figure 3.6. The tasks used in Experiment 2, from left to right: the line drawing, writing, and drawing.

In the first task, *line drawing*, participants drew a single vertical line, approximately 2 inches long, from the top to the bottom of the screen (Figure 3.6). Participants drew the line wherever they wished and were told to maintain the same length and speed across trials. Such a task was included because it required a short ballistic movement, had low cognitive load, and is commonly performed while annotating or sketching diagrams, (e.g., connect two boxes, underlining words, and so on). It thus provided a baseline against which the other tasks could be compared.

With the *writing* task, participants were instructed to write the word 'party' (Figure 3.6). 'Party' was used because it was required familiar, practiced movements and included characters that contained ascending and descending elements with a variety of curved and straight line components (e.g., 'P', 't', 'y'). Although participants may have been able to make a latency judgment after making a single stroke or writing a single character, they were required to write the whole word on every trial. They were also encouraged to use whichever writing style they were most comfortable with (i.e., printing, handwriting, or a hybrid of the two) and were told to write each character at whichever size they wished, but to maintain approximately the same character size across all trials.

In the *drawing* task, participants drew a six-sided star using one continuous stroke (Figure 3.6). A six-sided star was used because it contained varying angles, and was less automatic and familiar than other simple shapes. The increased attention and cognitive loading naturally encouraged slower, deliberate movements. Participants were encouraged to start drawing the star at the same location and maintain the same size of star and general shape across all trials.

All three tasks were counterbalanced to reduce any possible effects of learning and fatigue. A 1-pixel wide line displayed 'ink' while participants performed each task.

3.5.3 Results

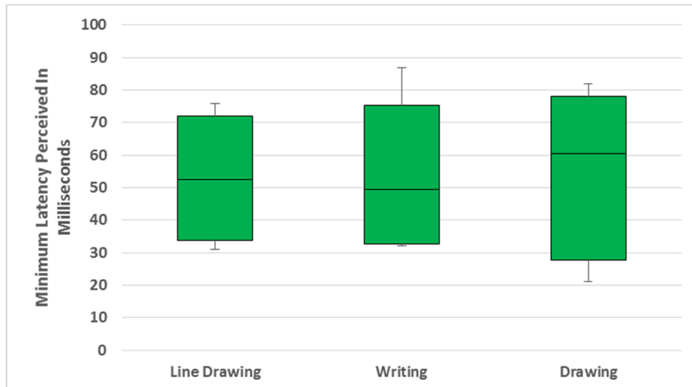


Figure 3.7. The median latency perceived by participants while performing the writing, drawing, and line drawing tasks. The box boundaries represent the 5th and 95th percentiles.

As the JND thresholds were not normally distributed, a non-parametric Friedman's ANOVA was performed with Task (i.e., line drawing, writing, and drawing) as the only factor (Figure 3.7). The ANOVA revealed that Task did not have a significant effect on participant's ability to perceive latency ($X^2(2) = 0.8$, $p = 0.667$, $\omega^2 = 0$). Participants were able to distinguish 7 and 53 milliseconds while line drawing, 7 versus 50 milliseconds while writing the word 'party', and 7 versus 61 milliseconds while drawing the six-sided star. The median latency across all tasks was 53 milliseconds. The lack of significance between the tasks does not suggest that perceived latency was, or will be, identical for all stylus-based inking activities. Rather, it suggests that other factors such as the visual cues and reference points available, or the motor movements required, may be more influential while perceiving latency.

3.5.4 Discussion

While performing everyday inking tasks, participants had a higher threshold for detecting latency than that found in the first experiment. For touch and simple box moving tasks, prior recommended target latencies were below 10 milliseconds (Ng et al., 2012). While performing inking tasks at the 50-millisecond level, many participants had difficulty distinguishing between the baseline and testing latencies, believing that they were the same, "are you sure these aren't the same, they look identical to me", "I swear most of these are the same", and "oh these are impossible now". While moving a simple box to and fro, or tapping on the screen, there is very little cognitive or attentional demands placed on the user, hence the lower perceptible latencies attained.

Even though sketching a shape and writing a word appear to be simple, familiar tasks, on the cognitive and attentional levels, they are much more complex. The slight increase in perceptible latency found between the inking tasks used presently and those in Experiment 1 is likely due to the present tasks being much more cognitively demanding, requiring pre- and post-planning to ensure that all strokes and characters are well formed, intra-strokes are joined, inter-stroke spacing is appropriate, and higher-level components such as characters and corners were created in the correct order and at the correct time. Such demands were not present in the scribbling task used prior. This suggests that attention and task demands likely influence the perception and detection of latency. This echoes the results found within the computer music literature and requires further investigation, especially in scenarios involving external environmental stimuli and indirect interaction.

Although not significant, there appears to be a larger range in perceptual variance as task complexity increases. As writing was more cognitively taxing then drawing the single line, participants had many opportunities and possible points of reference to use when making their latency judgments. With the line drawing task, the short, ballistic nature of the required movements left little time and a smaller set of reference points to judge latency. In the drawing task, many participants reported that they could not focus on latency as much as they could while drawing the simple line or writing because they intently focused on drawing all six points of the star. As the six sided star was an uncommon shape to draw, the increased availability of reference points (compared to the line drawing task) and focus required (compared to writing) lead to larger variability in the thresholds obtained. This observed variability may become even more prolific given a larger experimental population or different experimental stimuli.

In addition to task, post-experiment comments suggested that the natural sensorimotor processes and resulting locus of attention influenced latency perception. When asked how latency judgments were made, participants reported using a variety of strategies:

- Fixated on the eventual end location of the stylus and waited for the ink to catch up
- Fixated on one region of interest and estimated the time between the stylus / hand moving through the area and the ink appearing in the area
- Performed a pursuit movement, following the nib as it moved
- Performed a pursuit movement, following the ink as it appeared
- Alternated between the nib and ink (no saccades)
- Attended to the propagation of the ink's projected light through the translucent nib

Participants largely reported that depending on the task, they felt it was necessary to attend to different areas of the screen or visual cues. A graphic designer indicated that she focused on the global picture while inking, "intently focusing on the ink drawing the last few contour lines, not the lost, implied, or construction lines … the contour lines are the most important". Another participant commented, "when I take notes during a meeting, I rarely look at my tablet … instead I look at the speaker or their presentation. I only look at my tablet to see if I need to scroll for more paper, to fix a mistake, or to occasionally check that the pen is working". Based on such comments, it is clear that the strategy and location of focus are also implicated in the detection of latency.

3.6 Experiment 3: Influence of Locus of Attention and Visual References on Latency Perception

Inspired by the differing judgment strategies reported in Experiment 2, another experiment was conducted to determine the extent that the visual and motor systems work together to aid in the perception of latency. This third experiment explicitly manipulated the location of the digital ink, forcing attention away from the stylus into other areas of the screen, similar to a poorly calibrated stylus system. If participants focus on the relationship between the nib and ink to make latency judgments, offsetting the location of the ink a variety of distances, should decrease latency perception. If such information is not used, perceived latency should remain unchanged.

The effect of eliminating information regarding the motion of the stylus and hand, similar to indirect input scenarios was also of interest. In such scenarios, interaction and visual attention is naturally decoupled, distributed along different planes or devices. If latency is determined largely by the visual system, removing this reference from the visual field should impact the perception of latency. If latency is largely determined by other systems, such as audio or tactile feedback, signals from the motor system, or cognitive cues, latency perception should remain constant. The presence / absence of the hand within the visual field was also manipulated to mimic direct and indirect interaction scenarios.

3.6.1 Participants

Twelve naïve, right handed individuals (5 female) participated in the study (M = 33 years, SD = 7 years, range = 24-44). Similar to the first experiment, all participants had normal or corrected-to-normal vision and a range of experience with tablets and styli, some being experts and others complete novices. Thirty-three percent of participants were familiar with latency from playing video games or interacting with virtual environments. Participants were provided a $10 honorarium for the 30-minute experiment. None of the participants from Experiment 1 or 2 participated in this study.

3.6.2 Tasks

Four variations of the line drawing task from the first experiment were used in this experiment. The line drawing task was chosen over the writing and drawing tasks as it was the simplest, required less time to complete, and induced the least fatigue, all of which were important given the number of conditions.

In the first condition, *no offset*, the ink appeared directly underneath the nib (Figure 3.8). This was identical to the first experiment and enabled the location of the stylus nib, stylus barrel, and hand to remain in the foveal region. In the second condition, *small offset*, the ink was offset 6.5 millimeters, or approximately one index finger width, to the left of the nib. This offset diverted attention towards the ink, forcing the nib, stylus, and hand into the parafoveal region, closely mimicking scenarios where the stylus is inaccurate. In the third condition, *large offset*, the ink was further offset to the left, approximately 65 millimeters. This condition moved the nib, stylus barrel, and hand from the parafoveal region to the periphery and required much larger saccades to see both locations. Although the third condition would likely not exist in the real world, it was included to allow for a comparison with the last condition.

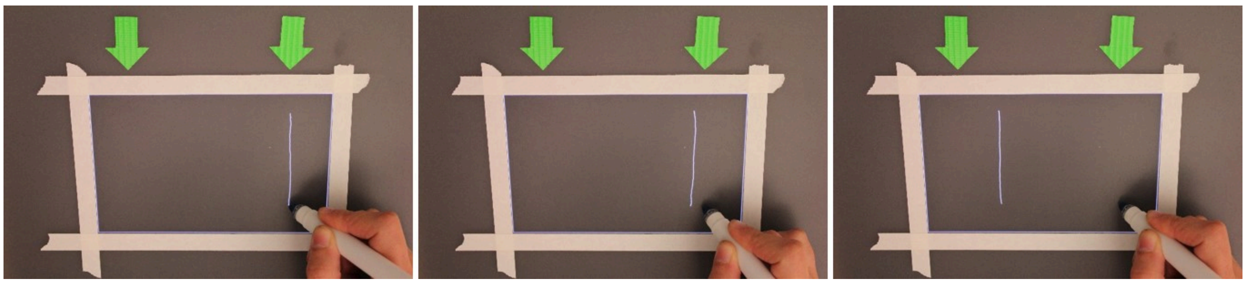


Figure 3.8. The offsets used in Experiment 2, from left to right: no offset, the small offset (i.e., 6 millimeters), and the large offset (i.e., 65 millimeters).

In the last condition, *hand not visible*, the ink was again offset 65 millimetres to the left of the nib but this time, the hand was additionally obscured from view using a foam board flange placed vertically in the center of the screen (Figure 3.9). Participants were instructed to attend to the left side of the screen, where they could only see the ink, not their hand. Participants did not receive visual information corresponding to the movements they were making.

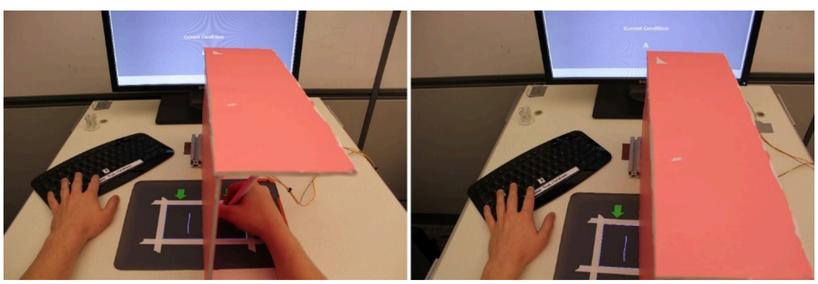


Figure 3.9. The hand not-visible condition, wherein participants placed their right arm and hand underneath the pink flange (left). A view of the experiment from the perspective of the participant, wherein they could not see their hand or the stylus (right).

Similar to the first two experiments, all four tasks were counterbalanced to reduce learning and fatigue effects. A 1-pixel wide line provided visual feedback and the experimenter controlled for the speed of drawing.

3.6.3 Results

As the threshold distributions were not normal, non-parametric analyses were conducted. A Wilcoxon signed-rank test evaluated the influence of viewing the hand with Hand Visible as the main factor (i.e., visible versus not visible). The results indicated that participants were able to better perceive latency when they could view the pen-wielding hand and stylus (*Mdn* = 59 milliseconds, range: 33-104) compared to not receiving this visual feedback (*Mdn* = 97 milliseconds, range: 59-105), $z = -2.9$, $p < 0.01$, $r = -0.58$ (Figure 3.10). The presence or absence of the hand and stylus thus appears to be an important referent and influences the perception of latency.

A Friedman's ANOVA examined the influence of the Ink Offset, (i.e., no offset, small offset, large offset). The results determined that Ink Offset did not have a significant effect on participant's ability to perceive latency ($X^2(2) = 4.2$, $p = 0.125$, $\omega^2 = 0.007$). Participants could distinguish 7 versus 59 milliseconds when no ink offset was present, 7 versus 50 milliseconds when a small offset was present, and 7 versus 59 milliseconds when a large offset was used. Irrespective of the offset used, the median perceived latency was 55 milliseconds, similar to that found in Experiment 2. As there was not a significant difference between the offset conditions, participants likely did not use the distance between the nib and ink to make their judgments, instead relying on the relative movement of the stylus or hand.

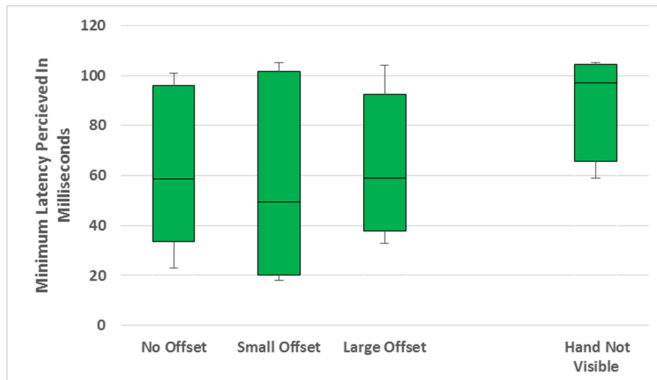


Figure 3.10. The median perceived latency while the ink was not offset, offset a small amount (i.e., 6.5 millimeters), offset a large amount (i.e., 65 millimeters), and while the hand was not visible. The box boundaries represent the 5th and 95th percentiles.

3.6.4 Discussion

The lack of significance between the no, small, and large offset conditions suggests that the distance between the stylus and digital ink is little help while perceiving latency. Many researchers believe the gap between nib and ink is used for judging latency, but participant comments and lab-based studies from the eye-tracking literature corroborate with the current results and suggest the opposite. It is actually uncommon for participants to follow the nib continually with their eyes. Few participants reported explicitly focused on the nib, stylus, or distance between the nib and stylus in Experiment 2. This is surprising, given that one would assume such visual elements would be the first cues users would look for, as the stylus initiates interaction and the ink provides feedback about the action. Scanpath analyses, however, have found that whenever the nib is located in the parafoveal or foveal region, it is often not attended to (Alamargot et al., 2006; Coen-Calgi et al., 2009; Fayol, Alamargot, & Berniner, 2012; Miall, Imamizu, & Miyauchi, 2000; Tchalenko & Maill, 2009). Eye-movement patterns have also been found to be largely task and motivation dependent, some preferable for quick inking movements such as sketching, whereas others are more appropriate for reading or editing (Gowen & Miall, 2006; Tchalenko 2007; Tchalenko & Miall 2009; Toyoda, Yamamoto, & Yubuta, 2011, 2012). Future work is thus needed to understand the specific visual details important for latency perception and where attention is directed during natural inking tasks.

The significant differences found between the hand visible and not visible conditions, in addition to work from the eye-tracking literature, suggests that the motion of the larger elements such as the hand or stylus barrel are valuable cues. In the motor-only condition, i.e., hand not visible, performance plummeted because participants were unable to rely solely on the haptic feedback from their pen-wielding hand or the signals from their motor system to make latency judgments. Once the stylus and hand were visible, even if only in the periphery (i.e., large offset condition), they provided valuable information to participants, in the form of a large moving stimulus. When such visual cues are not present, participants are forced to use cues from other modalities (e.g., haptic, audio, proprioceptive, or cognitive), so performance suffers.

To notice the latency inherent on stylus-enabled devices, it appears necessary for users to see their own hand in their field of view. As interaction and attention were visually and physically divided during the no-hand condition, the increased latency threshold observed from the large offset (59 milliseconds) to motor-only conditions (97 milliseconds) participants may have more difficulty perceiving latency on indirect input devices where the movement of the hand is out of view of the display. Although traditional indirect input devices separate input and output along different planes or devices, the motor-only condition mimicked such a scenario quite well. Such findings have implications for the future design and continued use of stylus devices that harness indirect interaction, such as the Wacom Bamboo Connect or Intuos devices. On such devices, sub-100 millisecond latency may not be required. A more focused study would be needed to examine other factors involved with real-world use of indirect input devices (e.g., placement relative to screen and user, size of input space, and so on).

Although there was no difference between the various offset conditions, a few users commented that they preferred the small offset condition because it was "similar to those signature pads at Home Depot or Lowes where the ink is far away from the pen location" and "allowed me to focus on the ink and still see the pen nib without having the nib occlude things or get in the way". Such comments suggest that a pixel-perfect calibration and accuracy may not be needed for a satisfying stylus experience.

Participants additionally commented that varying the speed of their strokes helped them perceive latency, but this strategy is not supported by the results. While increased pen speeds will increase the visible gap between the pen nib and the visual ink trail, in theory, this will make it easier for participants to perceive lower latency levels. Based on the results from the third experiment, this does not seem to be the case. If participants were perceiving latency based on the distance between the nib and ink, as predicted by Weber's law (i.e., the just-noticeable difference between two stimuli is proportional to the magnitude of the stimuli), performance should have decreased as the offset increased.

3.7 Overall Discussion

A great deal of attention has been devoted towards latency perception recently, but there still is a great deal open for exploration. Miller's (1968) 100 millisecond latency hypothesis focused on the issue of latency tolerance whereas the present experimentation examined perceptual thresholds. There is of course a difference between what users can perceive and what they will tolerate. Although we cannot provide tolerance recommendations, if users are unable to perceive delays below a certain threshold, then it is likely that they will tolerate delays at or near these thresholds. It is equally likely that they may tolerate much higher latencies as Miller predicted. Understanding the relationship between perception and tolerance thus remains an important, fruitful area of research.

This work explicitly focused on the display aspect of latency, manipulating the speed at which input was rendered. Across the experiments, participants knew that regardless of how fast or slow they interacted, all strokes would be sensed by the system and eventually appear on screen. With current devices, it is often unclear why some input is not sensed or displayed. The experiments determined the display latency that can be perceived but not the effects of delayed or slow sampling. While decreasing latency along the whole pipeline is advocated, it remains to be seen how perception and user satisfaction would change if devices rendered quickly but stroke completeness and accuracy were unpredictable.

3.7.1 The Latency Perception Model

While latency is simply the "delay between input action and the output response" (Mackenzie & Ware, 1993), the previous and current explorations into latency have determined that the perception of latency is a complex, multi-faceted problem. In Experiment 3, the input action and output response remained the same, yet the perception of latency changed. Initially, participants could use visual information to make judgments but once that was removed, they were forced to use other information streams, perhaps auditory or tactile cues from the stylus. These alternative data sources affected latency perception. Based on the present work and the prior literature, a model that describes the perceptual processes underlying latency perception in stylus and touch interaction has been developed. The model is composed of five elements: an input *action*, a *referent* stimulus, a *latency source*, output *responses*, and *contextual demands* (Figure 3.11).

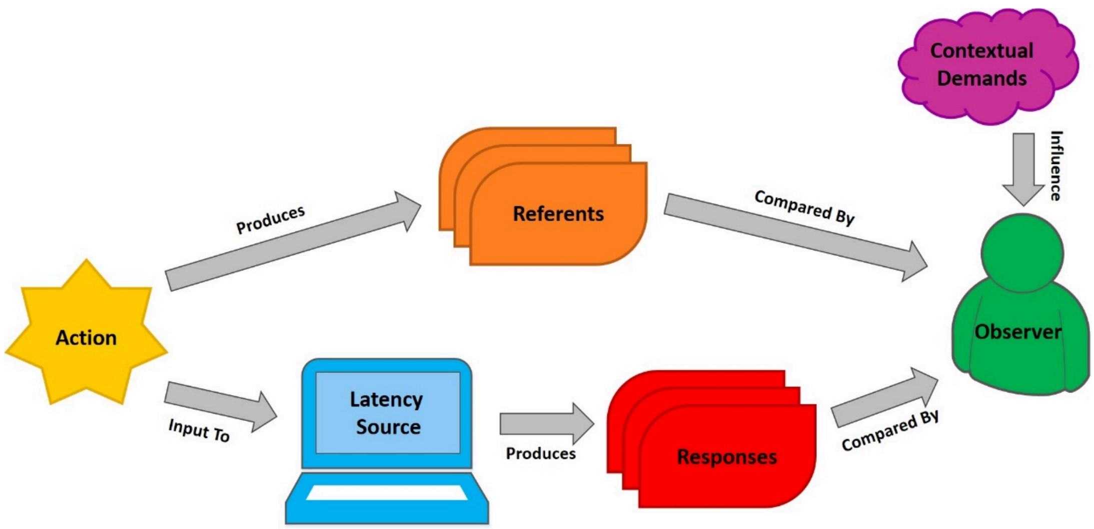


Figure 3.11. The Latency Perception Model, detailing the role of the input action, the resulting referent stimuli, the latency source, and the output responses. The observer compares the referents and responses when perceiving latency and is likely influenced by contextual demands such as task requirements, loci of attention, environmental factors, and so on.

The *input action* can take many forms, e.g., hovering the stylus in the air, touching a finger to the screen, or pressing the stylus against the screen, and is invoked by the observer, another user, or an external system or device. Once the action occurs, it is handled by the *latency source* (e.g., a sensor array, operating system, application, and so on). This entity converts the input action into output *responses* and adds delay. These responses are most often visual (e.g., a dot, line, or simple shape), but could manifest themselves via other modalities (e.g., haptic or auditory) as well. As the experiments demonstrated, attentional changes to responses or variations in the location of responses can influence their use. The results from the first experiment suggest that the spatial magnitude of responses also affects latency perception.

In addition to supplying the latency source with input, the input action also generates a variety of stimuli, or *referents* that provide clues to the observer (e.g., stylus barrel, fingernail, hand, stylus nib). Different modalities of the referent likely influence perception as well. In prior work with tapping, the haptic sensation of the finger pad touching and moving along the surface produced an additional referent that assisted in the perception of latency (Ng et al., 2012; Jota et al., 2013). As the stylus naturally dampens the haptic sensations from the screen, it is likely that in stylus-based scenarios, haptic referents play less of a role. Similar to the output responses, there is spatial and temporal uncertainty about what influences the referents. The referent may also need not be a physical stimulus, but may take the form of a cognitive initiation of an action (i.e., a mental 'Go' signal).

Once the referents or responses are available, the observer compares the original input action to the referents and responses to determine the magnitude of latency generated by the latency source. During this comparison and the decision making process, the experimentation determined that there are a variety of *contextual demands* influencing the observer. As suggested by Experiment 1 and 2, the judgment strategy plays a role in focusing or diverting attention from referents and responses, as does the location and amount of attention and external environmental distractions. Although definitive comparisons cannot be made, tasks that require more attention and increased cognitive load (i.e., inking versus moving a box) seem to redirect resources away from perceiving the referents and responses, making latency judgments more difficult.

Although little is known about latency perception, the model naturally provides many avenues for future explorations. By isolating each factor in the model and examining its effects, it is possible to extend the model, such that one could predict just-noticeable difference thresholds when different referents and responses are available, without having to evaluate the role of each explicitly. The model also raises a number of questions. For example, does each modality have its own cost when judging latency? Is there a constant cost for having the referent and response in different modalities? What is the relative impact of referents versus responses? From a psychological and interaction perspective, it is imperative to understand the processes governing latency perception before recommendations for future systems are made.

3.7.2 Improvements Necessary for Low Latency In the Future

The High Performance Stylus System used for the experiments was not intended to be a commercially viable technology. Rather, it was a system through which extremely low latency stylus interaction could be experienced and evaluated. Bringing such experiences to future commercial systems will involve substantial development and innovation across all sensing, processing, and display subsystems.

Until latency is decreased across the entire pipeline, low-fidelity, high-frequency feedback should be provided. The Accelerated Touch toolkit has been proposed previously as a method to decrease perceived latency (Ng et al., 2012). By combining high performance and current generation hardware, one can render both low and high latency feedback and visualizations. Ng et al. hypothesized that the combination of two such 'layers' of visual information could be used to mitigate many of the issues and complaints about latency. While certainly possible today, such an approach does however require the augmentation of existing hardware with a low latency system.

There are also a number of software-based pen-specific enhancements that can be implemented today. The use of pen location and stroke prediction algorithms can not only increase the smoothness and beautification of strokes, but can also be used to pre-render strokes before they occur and then adapt them after the stylus has reached or surpasses a given point. Instead of taking the time to render a smoothed, high-quality line, for example, it might be fruitful to initially render a quick, crude stroke and later replacing it with a smoothed line when more processing is available may be fruitful to consider. Three-dimensional modelling programs already make use of such an approach, rendering a wireframe while 3D models are manipulated and a full mesh while models are static. The current application context and knowledge about hand postures or grip could also be useful in sampling sub-regions of the input sensor and redrawing targeted sub-regions of the display. Vogel and Balakrishnan's (2010b) hand occlusion model, for example, could be implemented within an inking application to define the region of interest. As the location of the stylus reveals much about the intentionality of current and future interaction, harnessing it within applications should help decrease the latency perceived by users and ultimately improve the stylus experience both today and in the future.

3.8 Summary

The experiments presented in this chapter dove further into understanding the basics of latency perception using a prototype stylus-based system, the High Performance Stylus System (HPSS). The HPSS was composed of two high-speed DLP projectors and a fiber optic stylus. Such a system enabled participants to experience latencies as low as approximately one millisecond while dragging objects and seven milliseconds while inking.

Through experimentation, it was determined that users can perceived latencies at substantially lower levels than those possible on devices today, i.e., 2 milliseconds while dragging a box and 40 milliseconds while scribbling. By offsetting the location between the ink and stylus or removing the hand and stylus from the visual field. The experiments determined that low-latency judgments are largely visual, using the relative movement of the hand and stylus to assist in the perception of latency, not the nib and ink locations. When the hand and stylus were not visible, participants were unable to distinguish latencies below 97 milliseconds. It was additionally found that task demands influence latency perception, with inking tasks degrading one's ability to perceive latency moreso than simpler tasks, such as dragging, used in prior work. These results informed the Latency Perception Model, a generalized theoretical model of latency perception that focuses on the role of referents, responses, and additional extraneous factors on the perception of latency. Such a model provides insight into the perception of latency and forms a foundation upon which future work can be undertaken.

Chapter 4

Unintended Touch

While annotating a document or writing notes it is common to rest the palm, wrist, or forearm on the screen for support, to reduce fatigue, and to steady one's page (Hancock & Booth, 2004; Siio & Tsujita, 2006)[3]. As demonstrated in Chapter 2, depending on the digitizer, operating system, and application, this natural behavior can unintentionally generate a variety of touch input that could invoke a gesture, manipulate content, or render markings on-screen (Annett et al., 2014a). When using such systems, however, there are many instances when the user would like the touch input to be recognized, e.g., while swiping to the next page or zooming in on an image. Such conflicting desires exemplify one of the major problems plaguing pen and touch interaction today: How does a system determine touch events that are intended, i.e., intentional interactions with the fingers, and those that are unintended, i.e., a by-product of the skin resting on, or grazing the display?

The need to accurately distinguish between intended and unintended touch input has been identified extensively in prior work (Annett et al., 2014a; Alcantara et al., 2013; Gerken et al., 2010; Hinckley et al. 2010; Lin et al. 2013; Matulic & Norrie 2012; Vogel & Balakrishnan, 2010a; Valderrama Bahamonde et al., 2013; Zelzenik et al., 2012), but has yet to be explored in-depth. Most developers and researchers today ignore the problem completely by turning touch off or apply existing work from occlusion-aware interfaces to the problem (Yoon et al., 2013). There is a lack of knowledge both in the literature and in practice regarding unintended touch and appropriate interventions.

Apart from drawing attention to the importance of understanding and solving unintended touch, the work in this chapter contributes to the understanding of the problem itself, detailing specific challenges developers and manufacturers are faced with when devising solutions to unintended touch. The current state of the art, in terms

[3] The contents of this chapter has previously been published as Annett, M., Gupta, A., Bischof, W.F. (2014). Exploring and Understanding Unintended Touch during Direct Pen Interaction. *ACM Transactions on Computer-Human Interaction*, 21(5).

of unintended touch algorithms that exist in industry and in research, is explored. Motivated by the diverse approaches that have been tried, an experiment that gathered motion capture data from a stylus and tablet, and raw data from a touch and stylus digitizer while participants were writing, drawing, and annotating documents is detailed. The dataset of unfiltered and unprocessed touch data was recorded directly from the digitizer and permitted an evaluation and comparison of novel and existing solutions to unintended touch, given both the technological capabilities today and those likely in the near future. A number of natural user behaviors that influenced unintended touch requirements and algorithm success are also detailed. The results from this exploration and the behavioral patterns identified have been synthesized into two broad areas for future work that should be of great interest to the pen computing community and stimulate the stylus ecosystem for many years to come.

4.1 Understanding Unintended Touch

Within industry, the problem of unintended touch is often termed, *palm rejection*. By its very description and name, palm rejection suggests the need to reject only the palm. As the forthcoming exploration will illustrate, users are likely to touch the display using their fingers, knuckles, palms, wrists, and forearms while inking and interacting. The presence and dispersion of touch data found across the screen suggests that 'palm rejection' must not focus exclusively on identifying and ignoring the palm. Researchers and industry practitioners should thus use more appropriately descriptive terminology, such as *unintended touch*, moving forward. Framing the problem as unintended touch, instead of palm rejection, inherently implies that all future approaches and conversations consider that all skin surfaces can unintentionally touch a device, not just the palm. This change in nomenclature will go a long way in drawing attention to the importance of the problem and lessen common misconceptions such as 'oh that's an easy problem – just reject everything larger than the finger'. As demonstrated, approaches based on these off-the-cuff comments are inadequate.

Determining if a touch event is intended or unintended is a multi-faceted problem influenced by many factors. These factors have been distilled into three main categories: the time and data available for decision making, the underlying digitizer and types of input that need to be disambiguated, and the cost of recovering from incorrect rejection.

4.1.1 Timing and Available Data

The rejection of unintended touch can be performed at three stages along the data pipeline (Hinckley & Wigdor, 2002): in the firmware, in the operating system, or in the application (Figure 4.1). Each stage has different time constraints and different data available to make rejection decisions.

Solutions to unintended touch at the firmware level have full access to raw, unfiltered data. The raw data is the least abstracted and most informative available along the pipeline and its usage calls for primitive rejection solutions involving input shape, simple spatial heuristics, or raw sensor magnitude values. Rejection performed at this stage does not allow per-user personalization or the integration of factors such as task context, as the information being passed through the digitizer's controllers to the operating system is already close to or at bus capacity. Developers can, however, ensure a consistent user experience across operating systems and devices with the same digitizers and firmware if rejection is performed at this stage. The caveat of course is that decisions need to be made very quickly so as to prevent introducing latency into the entire system. Through consultation with industry practitioners, hardware designers, and manufacturers, it was understood that there is less than 5 milliseconds allotted to make a rejection decision.

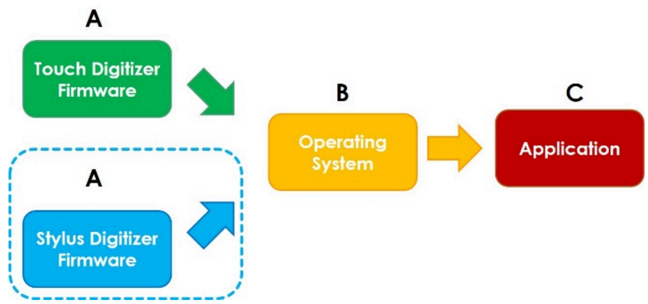


Figure 4.1. The data pipeline from the digitizer to an application. Rejection can be performed at three stages of the pipeline: (a) within the firmware of the device, (b) within the operating system, and (c) within individual applications. Depending on the device, the stylus digitizer may be absent (i.e., capacitive devices) or integrated with the touch digitizer.

Unintended touch algorithms at the operating system level can provide a consistent user experience across all applications on the same platform, but run the risk of inhibiting or eliminating certain interaction techniques or domain-specific functionality, such as simultaneous or interleaved pen and touch input (Hinckley et al., 2010a,b). At this level, solutions have access to processed and simplified data from the digitizer, and can harness user specific information such as handedness, hand posture, and so on. Time requirements at this stage are similar to those at the firmware stage, as any delay propagates to all applications. With the Windows 7 operating system, for example, all touch events are rejected when the pen is detected in the hover state (Buxton, 1990), resulting in consistent but restrictive pen and touch experiences across all applications. Similar to the firmware approaches, there is very little time allotted to make decisions, often less than 5 milliseconds.

Solutions to unintended touch at the application level can use personalized information about handedness or hand posture along with any domain or application-specific knowledge. Any touch information available to developers however, is abstracted: one does not have access to anything more than the location and radius or bounding box of a touch input that is provided by the operating system. The use of application-specific rejection also introduces inconsistency within the application ecosystem, resulting in confusion when rejection fails. By forcing the decision to the application stage, development costs for all applications increase, as unique solutions must be written for each application. The length of time available to make rejection decisions is limited only by the amount of latency developers want to inject into their applications. Each application thus weighs the cost of accuracy against the cost of adding increasing latency. As Chapter 3 demonstrated, latency perception appears to be task dependent, and depending on the demands of the application, more or less time may be available for rejection (Annett et al. 2014b).

When evaluating and understanding unintended touch, it is imperative to consider not only the accuracy of an approach or its data requirements, but also the duration of time available to make an 'accept' or 'reject' decision. Regardless of when rejection is performed, decisions need to be made very quickly. Participants from the user studies in Chapter 2 noticed and complained about the lag found with stylus-enabled devices today, hence any additional lag introduced by rejection algorithms will further degrade the user experience (Laundry, 2011; Ng et al., 2012; Ng et al., 2014; Annett et al., 2014b).

4.1.2 Input to Be Rejected

One of the most difficult aspects of unintended touch concerns the variety of touch inputs received from the digitizer. Many consider the problem restricted to the palm, as evidenced by the frequent terming of the problem as "palm rejection" (Zeleznick, 2012). On capacitive touch-only devices, however, any skin contact with the screen initiates touch input. Whenever the knuckles, wrist, palm, fingers, forearm, etc. touch the screen, a touch event is generated (Figure 4.2). As these devices do not have dedicated stylus digitizers, the stylus is undistinguishable from the fingers. This implies that the stylus first needs to be disambiguated from touch before each touch event can be labeled as either intended or unintended. With devices that contain NTrig or Wacom digitizers (e.g., Wacom Cintiq, Samsung Business Slates, and Surface Pro), distinguishing between pen and touch is not an issue, however, disambiguating between unintended and intended touch still remains. As the skin can touch the screen at any time and in any location, determining the intentionality of touch is still very difficult, even when an active stylus is used.

Although it may seem as if unintended touch is a "problem" that needs to be solved, it is also important to consider the identification of the palm, fingers, knuckles, forearm, etc. as an opportunity for novel interaction, gesturing, and so on. In earlier work by Cao et al. (2008), Marquardt et al. (2011), Widgor et al. (2011), Wu & Balakrishnan (2003), and Yan et al. (2013), the classification of a touch input as the palm, a fist, or the side of the hand, was harnessed to create novel gestures for 2D and 3D content manipulation. Brandl et al. (2009), used palm identification to correctly orient occlusion-aware user interaction menus and elements toward the user. In other work such information was instrumental in recovering the identity of users and their touch points, while they interacted around a multi-touch tabletop (Annett et al., 2011; Murugappa et al., 2012; Ramakers et al., 2012). One could also imagine many other scenarios that could capitalize on the identification of various hand parts, e.g., multi-pen input, technology-assisted rehabilitation, content occlusion, and so on. Given the multitude of possibilities outside unintended touch, we thus encourage the research community to continue exhausting and exploring the usefulness of hand-part identification.

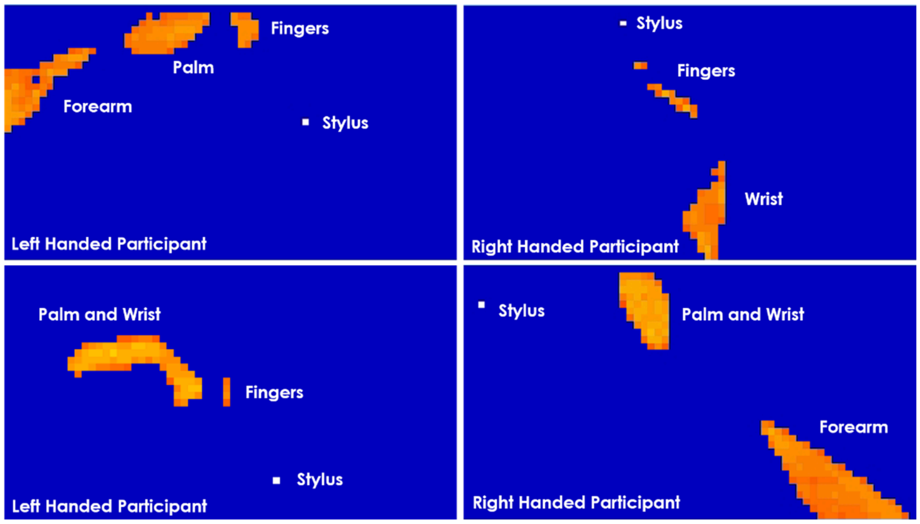


Figure 4.2. Example data provided by pen and touch digitizers. The stylus location is represented in white and the touch data is represented in orange, with the darker shades representing the higher levels of activation. All touch data in these images is unintended and should not generate touch events.

4.1.3 Cost of Recovery from Incorrect Rejection

It is imperative to consider the cost of imperfect rejection. There is a delicate balance between being overly cautious, i.e., falsely rejecting too many intended touch events, and overly optimistic, i.e., falsely accepting too many unintended touch events. If the user can easily recover from the errors introduced by poor rejection, e.g., when accidental touch causes a document to scroll prematurely, there is an easily accessible undo button, or if accidental ink appears, the user flips the stylus over and use the eraser, then it may be beneficial to be optimistic about touch events. Alternatively, if the cost of recovering from errors is high, e.g., the user has to switch modes to erase a stroke precisely or they have to undo a series of accidental manipulations, then a cautious approach may be best. It is thus essential for developers and designers to consider the cost of recovering from errors when deciding about the functionality associated with pen and touch and when designing unintended touch solutions.

4.2 Current State of the Art

Current approaches to unintended touch are commonly spurred by the desire of application developers to implement solutions that mesh with their end-user population and interaction schemas. In instances where solutions have yet to be implemented, there are several user-specific behaviors that have been employed to fill the void. The following review of current solutions to unintended touch is not meant to be exhaustive, given that numerous applications are continually updated with novel or improved solutions every week. Instead, an overview of the range of possibilities that have been used is provided.

4.2.1 User Adaptations

When touch rejection algorithms are unavailable in applications, users tend to make behavioral adaptations. By holding the hand, wrist, or forearm aloft instead of resting them on the screen, users can reduce the amount of incidental contact with the screen (Chapter 2; Gerken et al., 2010; Vogel & Balakrishnan, 2010a; Annett et al., 2014a). Since the lower arm is not stabilized, users experience increased fatigue, decreased pointing accuracy (Matulic & Norrie, 2012), decreased legibility of notes, diagrams, or sketches, and increased frustration (Chapter 2; Annett et al., 2014a). Some users have resorted to wearing a non-conductive glove on their dominant hand to prevent unintended touch (SmudgeGuard, n.d.). The use of such gloves negates the need to alter one's hand or wrist posture but changes the friction profile of the palm against the screen and adds overhead to using the device as the user has to put on and take off the glove between uses.

4.2.2 Firmware Approaches

Recent advances in signal processing have enabled manufacturers to better disambiguate between pen and touch. By increasing signal to noise ratios, Atmel (n.d.), Synaptics (Coldewey 2011; Sage 2011), Samsung (SmartKeitai, 2013) and many others have made it easier to differentiate between narrow activation signals generated by thin, pointed objects such as styli, and larger, bulbous objects such as fingers, palms, wrists, or forearms. The presented prototypes have eliminated disambiguation on capacitive touchscreens but have yet to solve unintended touch, as the intention of touch-based events is unknown.

It is difficult to determine if any firmware-specific approaches have been implemented, as the average developer only has access to abstracted data available from the operating system. It has been reported, however, that the Microsoft Surface Pro has an unintended touch algorithm in the touch and stylus controller firmware, i.e., 'Palm Block Technology', but few details are available regarding the nature of the implementation (Microsoft, 2013). Regardless of the algorithms employed by the Surface Pro, the resulting data is propagated to the operating system and combined with the additional unintended touch approaches found in the Windows operating system, as well as any approaches that specific applications may use.

4.2.3 Operating System Approaches

To date, there is one approach to unintended touch that has been implemented at the operating system stage. On Windows 7 and 8 devices, whenever the stylus is in the hover state (Buxton, 1990), all touch events are considered unintended and rejected. This prevents simultaneous interaction from occurring whenever the stylus is within range. Such an approach requires a digitizer that detects hover (or a Bluetooth-enabled stylus reporting hover), but is challenged by the non-linearity of the digitizer and variety of hover heights found on devices. In Windows 8.1, for example, a hover height of 20 millimeters is required for logo certification (Microsoft, 2014), but even at this height, many incidental contacts are still processed. The present evaluation takes these limitations into account and determines the hover height necessary for this approach to be successful from the behavioral and digitizer perspectives.

4.2.4 Application-Specific Approaches

Most unintended touches are generated by areas of the skin that are substantially larger than the fingers, e.g., the forearm or the palm. Applications such as the Bamboo Paper on the iPad capitalize on the different sized areas that the skin activates on the digitizer and ignores areas above a certain threshold, similar to what Murugappan and colleagues (2012) proposed for multi-touch tabletops. While simple, such an approach requires that input processing must be delayed because all contacts initially appear very small on the digitizer, making it virtually impossible to distinguish the palm from the finger on initial contact. If some latency is acceptable, this approach can work well, but fails if a finger-sized contact, e.g., the knuckle or fingertip, unintentionally touches the screen. The present analysis focuses specifically on the usefulness of contact size for rejection, identifying the appropriate threshold for rejection and understanding the type of touch inputs and use cases that could cause such an approach to fail.

When access to raw touch data is not available, developers sometimes employ simple models based on where the palm is hypothesized to rest while taking notes, e.g., the bottom of the screen. These static 'rejection regions', i.e., areas of the screen where all touch input is ignored, are used to reject unintended touch quickly and with little processing overhead. In the Moleskin Journal iPad application, for example, a static horizontal widget persists in the bottom third of the screen, rejecting all touch input. In Samsung's S Note application and in OneNote 2013, the rejection region is extended such that it covers the whole inking canvas but not the menu or toolbars. A dedicated menu button allows users to toggle touch input on and off, which allows the pen to ink exclusively. Shu (2013) proposed a variant of this approach, allowing the user to specify the location of the 'safe' input region by tapping near the bezel with the non-dominant hand and manipulating the region's size using a 'zoom in' gesture. Coupling a visual representation (e.g., a colored region or a menu button) with rejection allows users to understand where they can rest their palm, decreasing the likelihood of unintended touch. These approaches, however, prohibit efficient simultaneous and interleaved pen and touch interaction and cannot accommodate rotation of the hand or tablet while sketching (Fitzmaurice et al., 1999). As there are many ways to define and place rejection regions on the screen, the present work explores a variety of possible regions using static heuristics similar to those of the Moleskin application, as well as novel dynamic parameters such as the location of the stylus. This extensive evaluation leads to a deeper understanding of rejection regions and to solutions that harness generalized information. The design of feedback to alert users about safe and unsafe regions of the screen is left for future work.

Handedness information or models of common hand postures or touch input can also be used to improve rejection algorithms. Vogel et al. (2009; Vogel & Balakrishnan, 2010) proposed a general geometric model of hand occlusion as a method for eliminating problems associated with occluded screen content. In their model, the hand and wrist are generalized as a circle and rectangle whose location and size are based on aggregated video data from a sample of users. In Text Tearing, Yoon and colleagues (2013) have recently used the model to overcome issues with unintended touch (Yoon et al., 2013). A similar approach is used by the Penultimate iPad application, which requires users to specify which of three models of hand posture is most similar to their own. It then uses this information to reject all touch events that occur in a hand-like shape to the right (or left) of the stylus. As Chapter 2 demonstrated, there are many possible hand postures and grips while inking and interacting with pen-enabled devices (Annett et al., 2014a, Song et al., 2011, Levy & Reid, 1978). The Vogel approach was included in the experiment as an example of the narrowest possible dynamic rejection region. Recently, Schwartz and colleagues (2014) proposed the use of a spatiotemporal approach that harnesses the radius of each touch event to determine if it should be rejected or not. With their approach, the size of each touch input was compared to a dataset and machine learning techniques were applied for rejection resulting in the acceptance of 98% of all touch events. While this seems impressive, such results come at the cost of delaying the ink strokes that were presented to users by upwards of 100 milliseconds. As previously identified, the introduction of any latency into the pipeline will be noticed by users and influence their behavior, making such an approach infeasible given strict requirements detailed in Section 4.1.1., Chapter 2, and Chapter 3.

4.3 Data Collection

To better understand natural skin contact with pen and touch enabled devices, a data collection experiment was conducted to gather motion capture and raw digitizer data while users were inking. This data provides information on natural inking behavior and forms a basis for the analysis of approaches to unintended touch.

4.3.1 Participants

Eighteen participants (3 female) were recruited from within Microsoft Research to participate in the study (*M* = 42 years, *SD* = 10 years, range = 25-61 years). Participants were naïve to the purpose of the experiment and were not aware that unintended touch was being evaluating to ensure that natural behavior, and not adaptations were being recorded. Four participants were left handed, as assessed by the Edinburgh Handiness Inventory (Oldfield, 1971). Participants had varied experience using a stylus with a tablet, with some using a stylus and tablet every day, others only a tablet, and a few having never used a stylus or tablet. All participants received a $10 honorarium at the conclusion of the approximately 30 minute experiment.

4.3.2 Experimental Apparatus

Participants were seated at a desk surrounded by a six-camera Optitrack motion capture system and three Microsoft LifeCam Cinema web cameras. The Optitrack system recorded the movement of the stylus around and above the tablet surface as well as the movement of the tablet itself. A custom C++ application recorded the location of the stylus and tablet at 100 Hz, with ±1 millimeter of error. The three web cameras were placed in front, to the right, and to the left of the interaction volume to record overt participant behavior. The cameras recorded the complete interaction volume with 1080p resolution at 30 frames per second (Figure 4.3). The iSpy open source software was used to synchronize, time stamp, and record the web cameras for later offline processing and analysis.

Participants were provided with a Sony Vaio Duo 11 tablet that utilized an NTrig touch and stylus digitizer. Plastic blocks of one centimeter height were affixed to the bottom of the tablet to ensure that the markers on the tablet would not be disturbed if the tablet was moved. During initial pilot testing of the hardware, software, and tasks, the blocks were not found to influence the size, shape, or location of touch and stylus input that was generated. Five motion capture markers were added to the tablet, allowing for its position and orientation to be recorded (Figure 4.4a, b). The tablet was oriented in landscape mode and participants were free to rotate or move the tablet but were instructed to keep it on the table and within the orange interaction area.

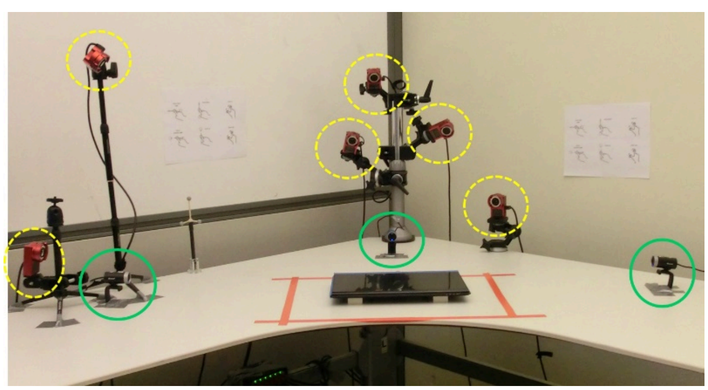
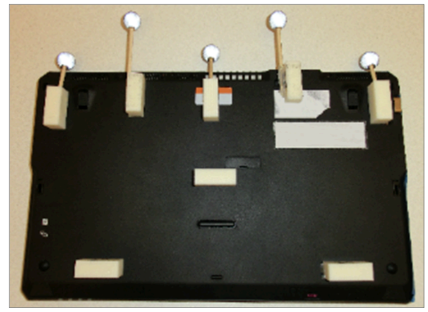


Figure 4.3. The experimental setup with the 6-camera Optitrack system (highlighted via yellow dashed circles) and three web cameras (highlighted via green solid circles). The interaction area was marked off with orange tape.

a. **b.**



c.

Figure 4.4. The motion capture markers attached to the Sony Vaio Duo 11 tablet (a, b) and the Ntrig stylus (c). Small blocks were affixed to the bottom of the tablet to ensure that the motion capture markers were not moved or damaged.

An active NTrig stylus was provided and participants were instructed to hold it in their dominant hand (Figure 4.4c). The stylus was modified by adding three markers on the back to track its position and orientation. For calibration, a fourth marker was added to the tip of the stylus, allowing the position of the stylus tip to be extrapolated from the remaining three markers during the experiment. The markers added 3 grams to the weight and 88 millimeters to the length of the stylus.

The tablet had a resolution of 1920 x 1080 and ran Windows 8. The tablet also ran a custom C# and WPF inking program that recorded all touch and stylus input from the 64 x 36 sensor array as raw, unfiltered antenna magnitudes. Each individual sensor detected a contact area of approximately 16 mm^2. The sensor magnitudes came directly from the digitizer via specialized software provided by NTrig and were converted into raw touch and stylus data. The data was recorded at 60 Hz and bypassed all Windows-specific formatting, filtering, and rejection approaches. Each ink stroke was rendered at 60 frames per second, due to the display refresh rate of the tablet's screen.

4.3.3 Tasks and Procedure

The experimental tasks were designed to capture realistic and natural inking behaviors. Each task was designed to encourage interaction in all areas of the screen, enabling digitizer data and behavior to be gathered from a variety of screen locations, not only the top or center of the screen. Three real-world inking tasks, i.e., writing, tracing, and document annotation, were designed for use in the study. As explained below, the annotation task was split into two sections, so in reality, participants completed four tasks. To prevent fatigue and learning effects, the presentation order of the activities was counterbalanced across participants.

Participants were instructed to perform each task at a speed comfortable to them and no time limits were imposed. Participants were instructed to interact naturally, as if using pen and paper. To facilitate natural behavior, they were informed that touch input was disabled so that they could rest their hand, fingers, wrist, forearm, knuckles, etc. on the screen if needed. They were also reassured that such behavior would not produce any stray markings or accidental gestures and that they would not see any feedback about any intentional gestures they were asked to make (e.g., swipes). While turning off touch input could influence the gestures that participants made, we deemed it necessary to prevent the digital-only hand postures that were observed previously in Chapter 2.

To focus the scope of the study, it was necessary to limit the number and type of tasks performed. Although activities such as web browsing with the stylus, dragging or repositioning content with the stylus, or using the stylus to tap on keyboard keys could have been examined, we wanted to begin the examination of unintended touch by focusing on the most common, and sought after, use case of styli, namely inking. As the chosen inking tasks required users to interact with all areas of the screen, the results should be somewhat generalizable to other scenarios as well. Future work could, of course, extend upon our exploration to identify and understand novel hand postures or usage patterns that are unique in these scenarios and explore how they would implicate solutions to unintended touch.

4.3.3.1 Writing

In the writing task, participants were presented with three 12-digit numbers and were instructed to write out each number in words (e.g., 123,456,789,101 had to be written as 'One hundred twenty three billion, four hundred…; Figure 4.5a). Participants were instructed to write on the ruled lines similar to traditional ruled loose-leaf paper and were encouraged to use whichever spacing and writing style they wished (e.g., printing, handwriting, mixed). No touch-based interaction was needed for the writing task, leading all touch events captured by the digitizer to be classified as unintended. The writing task took between five to ten minutes to complete.

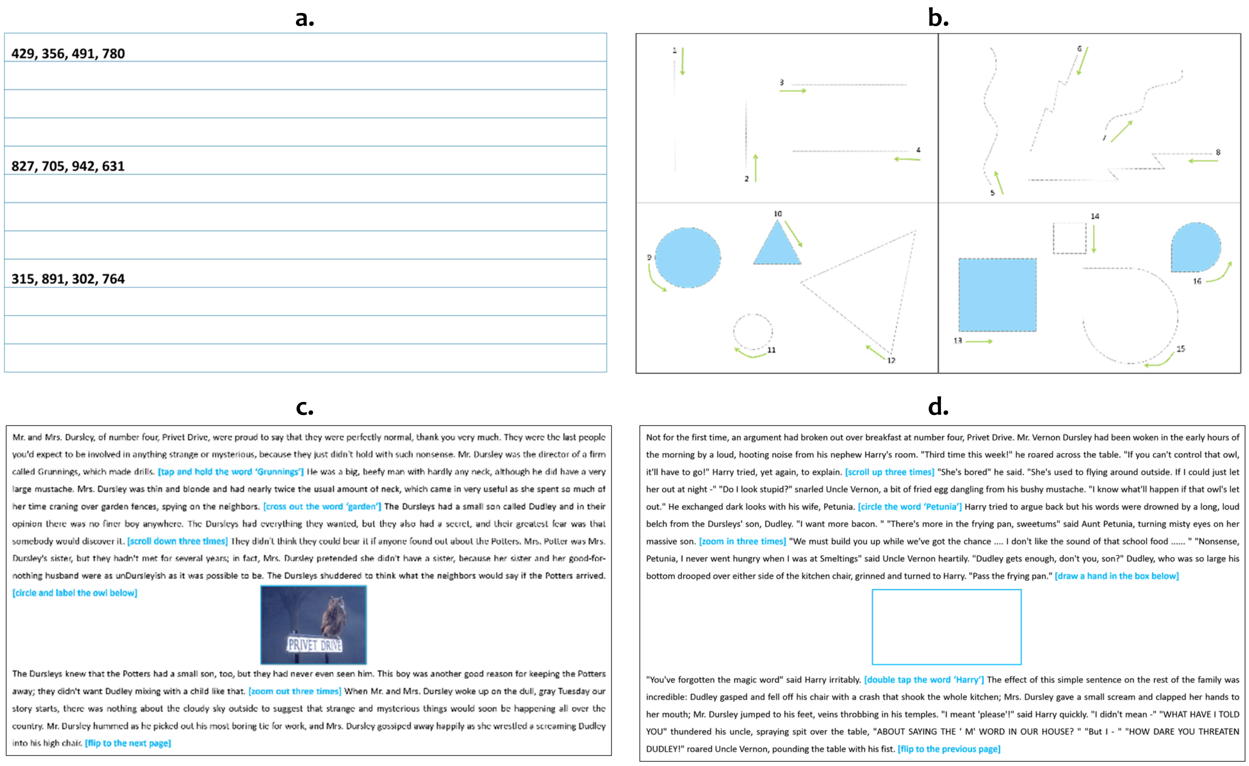


Figure 4.5. The tasks used in the data collection experiment: Writing (a), Tracing (b), first part of the Annotation task (c), and the second part of the Annotation task (d).

4.3.3.2 Tracing

During the tracing task, participants were presented with a number of dashed lines and shapes (Figure 4.5b). Participants were instructed to trace along each line or shape in the direction of the arrow, in order, starting with the vertical line marked ‘1’ and ending with the teardrop marked ‘16’. Whenever participants encountered a shaded shape, they were required to shade it in before continuing. A tracing task was chosen in place of a free-form drawing or sketching task to ensure that a variety of strokes and movements were performed by the participants and the movements would be relatively constant across participants. Similar to the writing task, no touch-based interaction was required. The tracing task took approximately five minutes to complete.

4.3.3.3 Annotation

In the annotation task, participants were presented with a short excerpt from a novel (Figure 4.5c and d). Embedded within the excerpt were short instructions that participants were required to perform. These instructions closely mimicked the behavior and actions found while reading or annotating a document, including, ‘scroll down’, ‘zoom in’, ‘flip to the next page’, ‘cross out a word’, ‘circle and label’, ‘tap and hold’, ‘double tap’, and ‘draw a hand’. To prevent participants from performing each action in close succession to the previous one,

and as many actions have a symmetrical counterpart, e.g., zoom in and out, the annotation task was divided into two parts. Participants thus performed the annotation task twice, using a different excerpt and set of target behaviors each time (i.e., one task contained 'zoom in', 'flip to the next page', 'scroll up', and so on, and the other task contained 'zoom out', 'flip to the previous page', 'scroll down', and so on).
As touch was disabled across all tasks, no gesture recognition was implemented, nor did participants receive feedback once they made their gestures. Each annotation task lasted between five and ten minutes. The embedding of such actions within the context of a larger reading task captured behavior in as close to a real world situation as possible and ensured that the dataset contained opportunities for simultaneous pen and touch interaction.

4.3.4 Data Synchronization and Processing

The web cameras and Optitrack data were recorded on a desktop computer, and the stylus and touch data were recorded on the Sony tablet, thus data synchronization was necessary (Figure 4.6). The web camera and Optitrack data were synchronized using timestamps from the Windows clock, as they were on the same machine. The touch and stylus data were synchronized using a common OS-level clock for recording timestamps. As identical motion capture data was streamed to both the tablet and desktop and identical C++ clients were used to record the data, the streaming data (not the timestamps) was used to synchronize the data across both machines.

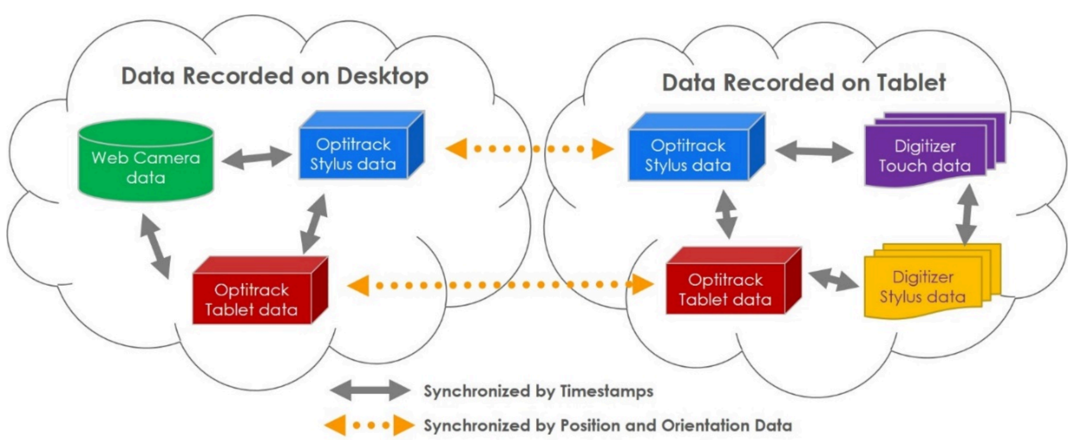


Figure 4.6. A schematic representation of the data sources that were recorded and the synchronization processes that were performed using the data collected by the experiment.

MATLAB was used to filter, process, resample, align, and synchronize the different data streams. Behavioral parameters such as the stylus height during inking and gesturing as well as the location and size of each touch event were also computed using MATLAB. The analysis and the comparison of unintended touch algorithms were performed using C#. When necessary, the VLC media application was used to analyze the video streams frame by frame.

4.4 Evaluation of Solutions to Unintended Touch

There are many factors and natural behaviors to take into consideration when designing approaches to unintended touch. Although several tablets report hover, pressure, tilt, and other sensor data, many do not. Approaches that make use of different sensor data and other information sources, such as hand models or user handedness, were thus analyzed.

4.4.1 Procedure

The ofxtouch OpenFrameworks (2009) connected-components algorithm was first applied to each frame of the collected data to group the activated sensors into ‘touch blobs’. Once the data had been segmented, each blob was tracked frame-to-frame and assigned an id using the ofxtouch blob tracking algorithm. In keeping with the desire to examine algorithms that introduced the smallest latency possible into the pipeline, the touch data from the first frame of the touch was used. This ensured that a decision would be made for every touch event, including those that lasted for only one frame. It also ensured that the speed of movement after the initial touch activation was not an influencing factor.

Although the location of the stylus was known for the duration of each task, for every frame of data collected, the motion capture data was used to determine whether the stylus was located on, above, or beside the screen. Whenever the stylus was not on, or directly above the screen, it was treated as being absent. This was necessary because algorithms that required the presence of the stylus or the precise stylus location would be provided with incorrect data (given the antenna-based techniques used in current stylus-based systems to detect the stylus). Given that four of participants were left handed, handedness-specific information was also integrated into the dataset to allow an analysis of those approaches that use spatially-based information (e.g., ignoring everything to the right of the stylus for right-handers and to the left for left-handers). Although displays today do not detect handedness, this was deemed necessary for an unbiased analysis.

To assess the different unintended touch solutions, each touch was first classified as either intended or unintended. For the writing and tracing tasks, every touch event was classified as unintended, since no touch interaction was required during these tasks. In other words, 100% of the 5085 touch events generated while writing and tracing were classified as unintentional and 0% were classified as intentional. For the annotation tasks, the only intentional touch events were those generated by the fingers while the participant was zooming, scrolling, flipping, etc. (see Section 4.3.3.3). To determine which touch events detected by the digitizer were intended or unintended, the web camera videos were consulted. As the order in which each participant performed each gesture was known a priori, the videos were manually analyzed to determine the timestamps corresponding to the beginning and ending of each gesture (i.e., 'the intentional touch time period'). These timestamps were then used to classify each touch event as intended (i.e., the input detected by the digitizer occurred during an intentional touch time period and the videos illustrated that it was generated by one or more fingers) or unintended (i.e., the input detected by the digitizer occurred outside the intentional touch time period or the input detected by the digitizer occurred during an intentional touch time period but the videos demonstrated that it was not generated the fingers). Of the 1218 touch events that were generated during the annotation task, 17% (i.e., 213) were classified as intentional and 83% (i.e., 1005) were classified as unintentional. Although the total number and proportion of touch events generated during the writing/tracing and annotation tasks are not equal, they are representative of the disproportionate nature of intentional versus unintentional touch while inking.

After the intentionality of each touch event was determined, a contingency table (i.e., confusion matrix) was computed for each task, participant, and algorithm (Table 4.1). In addition to being classified as intended or unintended, each touch event was further classified as either accepted or rejected. The acceptance or rejection of a given touch input was dependent on the parameters and techniques specified by the algorithm under assessment.

Table 4.1. The confusion matrix used for the unintended touch classification.

	Intended Touch	**Unintended Touch**
Algorithm Accepts Touch	True Positive (TP) or Hit	False Positive (FP) or False Alarm
Algorithm Rejects Touch	False Negative (FN) or Miss	True Negative (TN) or Correct Rejection

To determine the performance of various approaches to unintended touch, all of the cells in the contingency table, were used to compute the accuracy of each algorithm (Equation 1). A measure such as accuracy allows equal focus to be directed towards situations where the user intentionally interacted and the algorithm accepted the touch event (i.e., true positives) and situations where the user unintentionally interacted and the algorithm rejected the touch event (i.e., true negative). Although one may assume that the only important aspect of rejection should be the true positives, due to their necessity for bi-manual interaction, it is equally, if not more, important to consider how algorithms perform during tasks that do not require bi-manual interaction, such as writing or drawing, i.e., evaluate the true negatives. Accuracy allowed such a comparison.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100 \quad (1)$$

While one could use measures such as precision and recall, such measures discount the importance of the true negatives, i.e., unintended touch events that were correctly rejected. Given that users are more vocal about the stray inking and accidental navigation that is the result of poor rejection than their ability to perform gestures, it did not seem reasonable to use such measures. Different weightings could have also be assign to various cells of the contingency table, but we felt that this would have resulted in a biased analysis given the tasks users performed. Given the latency requirements detailed in Chapter 2, it would have also been appropriate to assess the running time of each algorithm. We opted not to perform this calculation, and instead focus on a measure that was hardware invariant and allowed the algorithms to be implemented at anywhere along the data pipeline.

4.4.2 Algorithmic Approaches

There are many possible approaches to unintended touch. In this work, six algorithms representative of the spectrum of possibilities today, and the foreseeable near future, were evaluated. The approaches were chosen based on their prevalence in existing work, hypothesized efficiency and dependence on additional data resources.

4.4.2.1 Reject All Touch Events

To set the context for the evaluation, the first approach that was evaluated was one in which all touch events, regardless of whether they were intended or unintended were rejected. This is similar to the approaches taken by S Note and OneNote, where all touch input is rejected by the application. Although simple, such an approach effectively eliminates all issues users have reported with stray markings accidental navigation, while simultaneously preventing gesturing or any other desired touch-based input.

4.4.2.2 Contact Area

As access to the raw digitizer data was available, the contact area of each touch event was computed. The contact area was defined as the number of sensors activated by each individual touch event, on the first frame that the touch event was detected by the digitizer. This approach to unintended touch compared the contact area of each touch event to a number of different thresholds and rejected those touch events that had an area bigger than the threshold. For example, if a touch event area had of 5 sensors and the threshold was 10 sensors, the event would be accepted; if a touch event area had of 16 sensors and the threshold was 10 sensors, the event would be rejected.

As a number of different thresholds are possible, the thresholds that were evaluated in the present study were in the range of zero to 2048 sensors (i.e., the maximum number of sensors on the device), in one-sensor increments (e.g., 0, 1, 2, 3, and so on). While some iPad applications make use of the diameter of a touch contact to make rejection decisions and it is often the first approach many think of when imagining how they would solve unintended touch, the use of raw sensor data for unintended touch is novel. Such an approach could easily be integrated within the firmware, operating system, or specific applications.

4.4.2.3 Hover-Based Rejection

This hover-based approach mimics the behavior of existing Windows 7 and 8 devices. With these devices, whenever the stylus is close enough to the screen that it can be detected (i.e., the stylus is in the hover state; Figure 4.7), all touch input on the screen is ignored. For the present evaluation, the Optitrack data was used to compute the distance between the screen and the stylus' nib for each touch input. Using this information, whenever the height of the stylus' nib fell below a pre-determined threshold, the whole touch input was rejected; if the nib height fell above the threshold, the touch input was accepted.

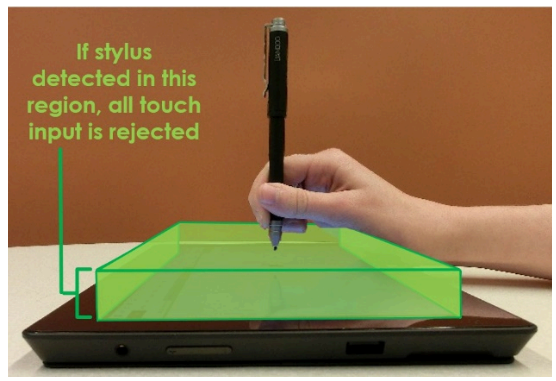


Figure 4.7. Example of the hover-based region wherein the all touch information would be rejected whenever the stylus is detected within the green region (box).

Current devices detect the presence of the stylus up to a height of approximately 20 millimeters, but, with advances in sensor technology, this could increase further. Such an approach harnesses the hover state and could make hover-based approaches a viable option because they allow rejection decisions to start well before the skin touches the screen. The use of a hover-based approach may also interfere with one's ability to gesture if the stylus is tucked under the finger(s) while gesturing. During such behavior, the height of the nib may still be within the detectable hover height and could prevent one from gesturing. It was thus of interest to evaluate this approach from a variety of possible detectable stylus heights (i.e., zero to 200 millimeters, in increments of 1 millimeter) to ensure that it would allow one to ink *and* gesture.

4.4.2.4 Hand Occlusion Model

A generalized version of Vogel et al.'s Hand Occlusion Model algorithm (2009) was also implemented. In this model, the hand and wrist were abstracted into a circle and intersecting rectangle and were located at a specified distance and angle from the stylus (Figure 4.8). All touch events falling within the circle and rectangle are rejected, whereas those outside these areas are accepted. With this approach, if the stylus is not detected, then any touch input detected by the digitizer is accepted. As from both right and left handed users participated in the data collection experiment, the angle and location of the 'palm circle' and 'forearm rectangle' were mirrored when the left-handed participants data was used.

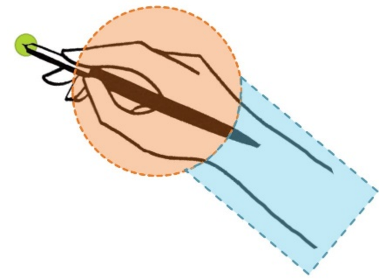


Figure 4.8. Abstracted illustration of the Vogel Hand Occlusion Model, wherein the stylus location (the solid green circle) is used to dictate the location of the abstracted hand (i.e., the dashed orange circle) and the forearm (i.e., the dashed blue rectangle).

The Hand Occlusion Model was evaluated because it has already been applied unintended touch in research (Yoon et al., 2013). The model is hypothesized to work well for unintended touch because it maximizes the area available for (bi-manual) interaction and ignores segments of the screen from where unintended touch is most likely, i.e., from the palm and forearm. In the present work, we decided to use the generalized model first proposed by Vogel (2009). Vogel's later work relied on the use of a calibration phase to generate personalized models for each user (2011). However, given the additional calibration steps necessary, developer's resistance towards such steps, and the desire to ensure participants were naïve to the exploration of unintended touch (i.e., a calibration phase would have alerted users to the underlying goal of the exploration), the generalized model was used instead.

4.4.2.5 Static Rejection Regions

Inspired by the approach used with the Moleskine application, the use of 'static' rejection regions was also evaluated. In these applications, rejection is performed by specifying a specific rectangular area of the screen that is 'safe' to rest one's palm or forearm (Figure 4.9). An overlay is typically used to make the 'safe' and 'unsafe' regions of the screen visually distinguishable for users and remains persistent until the user decides to close them. Rejection that is performed using these regions looks at every touch input that is generated on the screen, and compares the location of the touch input to the bounds of the 'safe' region. If the input falls inside the region, the input is rejected; if the input falls outside the region it is rejected.

Although Moleskin relegated their rejection region to cover the bottom third of the screen, there are many ways in which static rejection regions can be specified. Within the present investigation, three different region areas (i.e., filling 1/3 screen, 1/2 screen, and 2/3 screen; Figure 4.9) and three different spatial partitions (i.e., located vertically, horizontally, or having a horizontal-vertical intersection; Figure 4.9) were evaluated. The examination of the vertical and horizontal-vertical spatial partitions and the 1/2 screen and 2/3 screen areas are novel. The resulting nine rejection regions covered a range of possible areas, allowing for an extensive evaluation. As these approaches require only handedness information, rejection decisions can be made very quickly and along various segments of the data pipeline. For tablets that cannot distinguish between pen and touch or that do not support hover, static rejection regions could be a viable option.

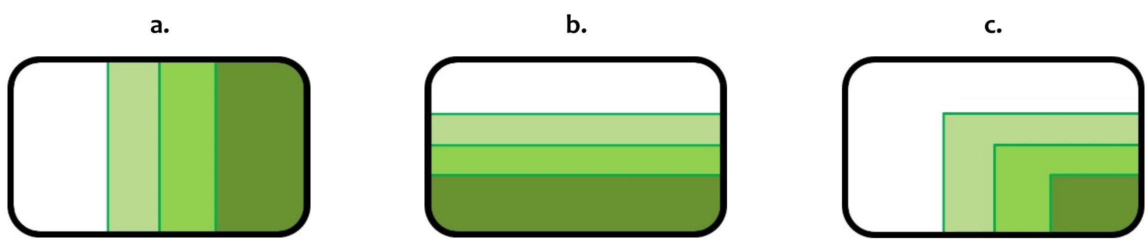


Figure 4.9. The nine static rejection regions evaluated for right handed users, with the area and location of the region being (a) vertical, (b) horizontal, or (c) intersecting horizontally and vertically. All touch events within the green areas were rejected. For left handed users, the vertical regions were mirrored horizontally.

4.4.2.6 Stylus-Based Rejection Regions

One of the downsides of the static rejection region approach is that the rejection region has little context or information about where one is currently interacting. As such, it cannot adapt to when one is inking in the top right corner (i.e., reject more of the screen) or in the bottom left corner (i.e., allow for intentional touch input everywhere else). A novel variant of the static rejection region approach that solves this issue is to use the current X,Y location of the stylus to provide context for rejection. As the X,Y position indicates where one is currently interacting, it can dictate the location and bounds of a rectangular rejection region (Figure 4.10). Consider the use of a vertical rejection region. If the stylus is located at x = 50 and the participant is right handed, a stylus-based rejection region would reject all touch events to the right of the stylus (i.e., x > 50) and accept those touch events that fall to the left of the stylus (i.e., x < 50). If the stylus were to move to x = 600, the rejection region would move as well, rejecting everything greater than 600 pixels and accepting everything less than 600 pixels. Similar to the Hand Occlusion model, if the stylus is not detected, all touch input is rejected.

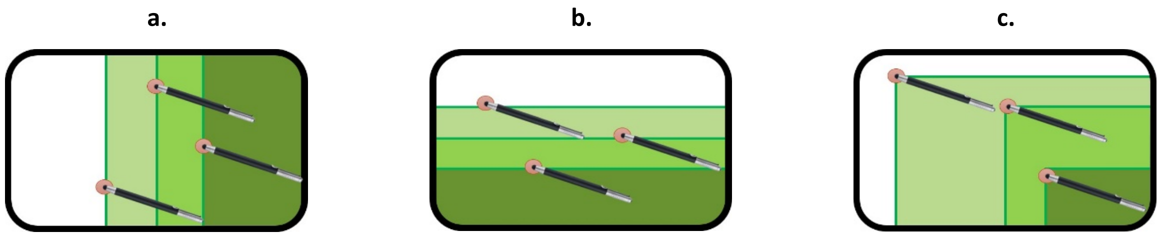


Figure 4.10. Examples of stylus-based rejection regions for right handed users, with the area and location of the region being (a) vertical, (b) horizontal, or (c) intersecting horizontally and vertically. All touch events within the green areas would be rejected. For left handed users, the vertical regions would be mirrored horizontally.

Similar to the static rejection regions, there are a number of spatial partitions that can be used to dictate the bounds of the rejection region. Within this exploration, stylus-based regions that were horizontal, vertical, or had a horizontal-vertical intersection were evaluated (Figure 4.10). Such regions should allow rejection to be performed quite quickly and along various segments of the data pipeline but does require the stylus location for the region to be specified.

4.4.2.7 Stylus-Based Rejection Regions Using a Buffer

Although the stylus-based rejection regions harnesses contextual information, they assume that the stylus will always be located above the palm, wrist, or forearm. As demonstrated in Chapter 2, not all hand postures follow this form. For example, hand postures that are inverted or 'hooked', result in a nib location below or to the right of the palm or fingers. Based on the variety of hand postures observed a novel variant of the stylus-based rejection region approach was evaluated. With this approach, a virtual 'buffer' was extended out and away from the stylus-specified region (Figure 4.11) to accommodate such hand postures. Now, any touch input located within the stylus-specified region or the buffer region would be rejected; touch input in any other areas of the screen would be accepted. If we consider the vertical screen partition, what would happen if a 75 pixel buffer (~10 cm) was used (and the user was right handed)? If the stylus was located at x = 875, for example, any touch input whose location was greater than x = 800 would be rejected, anything less than x = 800 would be accepted. Similar to the Hand Occlusion model and stylus-based rejection regions, if the stylus is not detected, all touch input would be rejected.

The use of 'buffers' should accommodate the variety of hand postures used while inking (Annett et al., 2014b) and should be more robust to wrist and hand rotations (Fitzmaurice et al., 1999). As the use of the stylus to specify the rejection region and 'buffers' are novel, a variety of buffer dimensions ranging from 0 to 50 millimeters were evaluated, in increments of 1 millimeter. Similar to the static and stylus-based rejection regions, spatial partitions that were horizontal, vertical, or had a horizontal-vertical intersection were evaluated. This approach should allow rejection to be performed quite quickly and along various segments of the data pipeline, but of course requires that the stylus location is known.

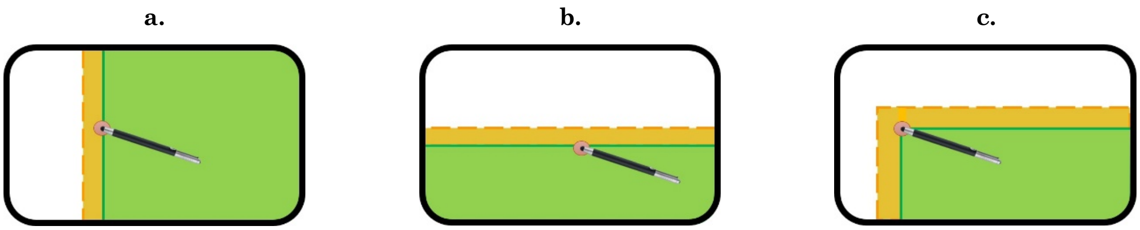


Figure 4.11. An example of stylus-based rejection with a buffer with the spatial partioning of the screen being (a) vertical, (b) horizontal, or (c) intersecting horizontally and vertically. In this approach, the stylus-specified region extends away from the stylus location towards the dashed orange line. Touch events falling into the green 'safe' region are rejected. For left handed users, the vertical regions would be mirrored horizontally.

4.4.3 Evaluation of Individual Algorithms

As the writing and tracing tasks did not contain any opportunities for intentional touch input, the proportion of intended versus unintended touch events within these segments of the data set is skewed. Collapsing across the four tasks and combining all of the task data together would have resulted in a biased accuracy measure. We thus chose to separate the tasks into those that contained opportunities for touch input (i.e., the two annotation tasks) and those that did not (i.e., writing and tracing). By evaluating these data sets separately, we were able to determining how an approach would fair in situations that are increasingly likely in the future (i.e., bi-manual input as represented by the annotation task) and those that users encounter today (i.e., writing and tracing without the desire for bi-manual input). For the analysis, behavioral data gathered from the webcams, motion capture system, and digitizer were used where appropriate.

As there are many possible parameter variations applicable for a given approach, e.g., 200 different hover heights or 2048 potential contact sizes, and each approach makes such of different data streams that were collected, e.g., hover height versus stylus location, others require both stylus and touch information, etc., we first report on an individual analysis of each approach. Section 4.4.3.8, *Aggregated Contingency Table*, provides a summary of each algorithm's best results, with respect to the raw contingency table values that were recorded. In Section 4.4.4, *Comparative Evaluation of Algorithms*, we identify the optimal parameters for each approach and utilize such information to statistically compare the algorithms against each other.

4.4.3.1 Reject all Touch Input

Unsurprisingly, taking an approach that rejects all touch input performs perfectly when no intentional touch input is generated, i.e., with 100% accuracy during the Writing and Tracing tasks (Figure 4.12). When the Annotation task is considered, the accuracy is proportional to the amount of unintentional touch input found in the dataset, i.e., 83%. These results provide context and a baseline against which the remaining six approaches, and their variations, are considered.

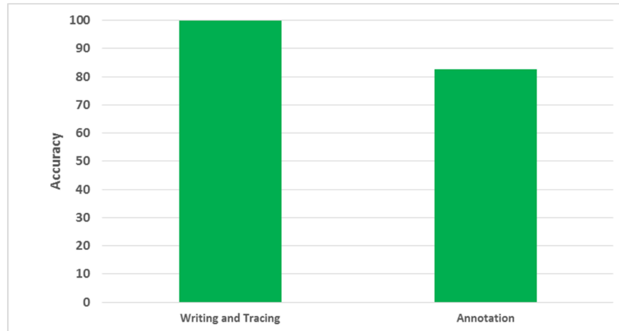


Figure 4.12. The accuracy results for the Writing/Tracing and Annotation tasks when all touch events are rejected.

4.4.3.2 Contact Area

The analysis determined that, regardless of opportunities afforded for intended touch, as the threshold of allowable contact areas increased, accuracy decreased (Figure 4.13a and b). Because the inter-sensor spacing of the digitizer was approximately 4 millimeters, the average finger activated a median of approximately four sensors (IQR = 3 sensors). Rejecting touch events at this threshold resulted in 51% accuracy while writing/tracing and 38% accuracy while annotating. Once the allowable contact area increased beyond four sensors, accuracy continued to decrease until it leveled off at approximately 60 sensors (i.e., 1% while writing/tracing and 18% accuracy while annotating). The disparity found when annotating versus writing and tracing emphasizes the small contact area that most touch events, not just the fingers, activate when they initially touch the screen.

a. **b.**

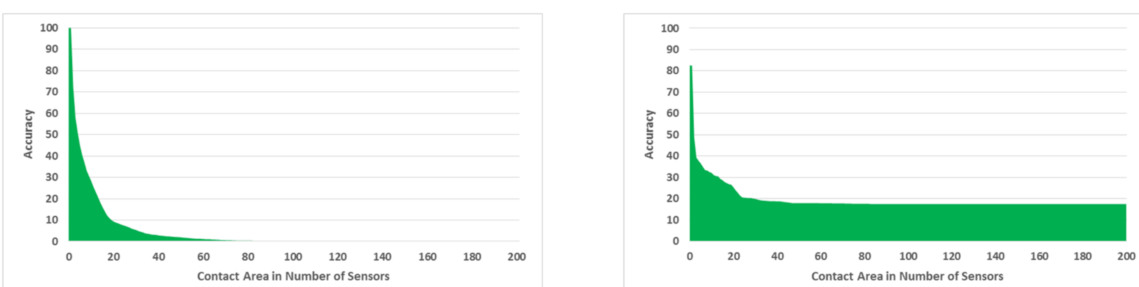


Figure 4.13. Accuracy results for the (a) Writing/Tracing tasks and (b) Annotation tasks. There were not changes in accuracy from 200 – 2048 sensors so these values have been removed from the graphs for clarity.

An analysis of the digitizer data makes it clear why the contact area approach is not sufficient. When we removed the intentional 'finger' touch events from the dataset, i.e., everything activating fewer than four sensors, and examining the initial area and growth of the remaining unintentional touch events, we found that most unintended touch events initially resembled a finger or group of fingers (Figure 4.14). If one looked exclusively at their size, intentional and unintentional touch events were thus indistinguishable. When we looked at the initial shape of the intentional versus unintentional touch events, they appeared visually indistinguishable. The activated sensors did not form ovals or circles, as one would expect (Murugappana et al., 2012). The activated sensors often appeared as straight lines or irregular, jagged shapes (Figure 4.14, at 0 milliseconds). Given the indistinguishable nature of these events, the use of contact shape or size thus appears to be infeasible as it is impossible to initially identify intentional finger-based touch input from other unintentional input.

We also looked at the growth of touch events over time to determine at which point an intentional touch event grows to be distinguishable from an unintentional touch event. It was determined that touch events initially activated a very small number of sensors (Mdn = 6 sensors, IQR = 9) and did not grow to a differentiable size (i.e., larger than 4 sensors) until approximately 33 milliseconds after their initial activation (Figure 4.15). Such a slow rate of growth explains why a contact-based approach cannot be successful: both intended and unintended touch events initially appear the same to the digitizer and take much too long to become differentiable. Waiting for the contact areas to merge and stabilize before making a rejection decision is not feasible, especially when current system latencies are easily perceptible by users and less than 5 milliseconds are available for rejection decisions. Although the system recorded touch and stylus data from the digitizer at 60 Hz, increasing the sensor density or sampling rate would have no effect on the real-world speed at which the contact area grows, leaving the accuracy largely unchanged.

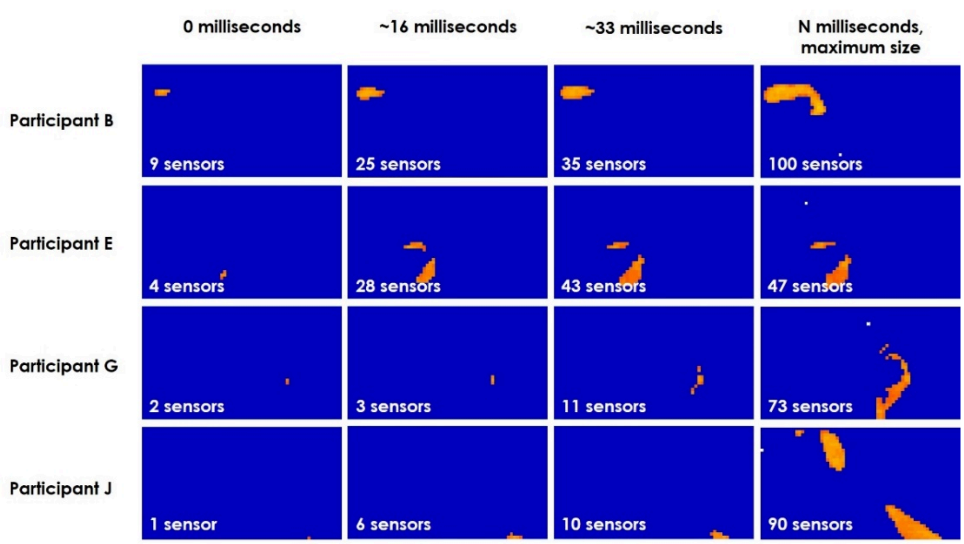


Figure 4.14. Examples of different touch events that initially activated a small number of sensors (denoted by shades of orange) and later grew to their full size. Note that the initial shape and direction of the touch input does not represent the eventual shape and orientation of the touch input.

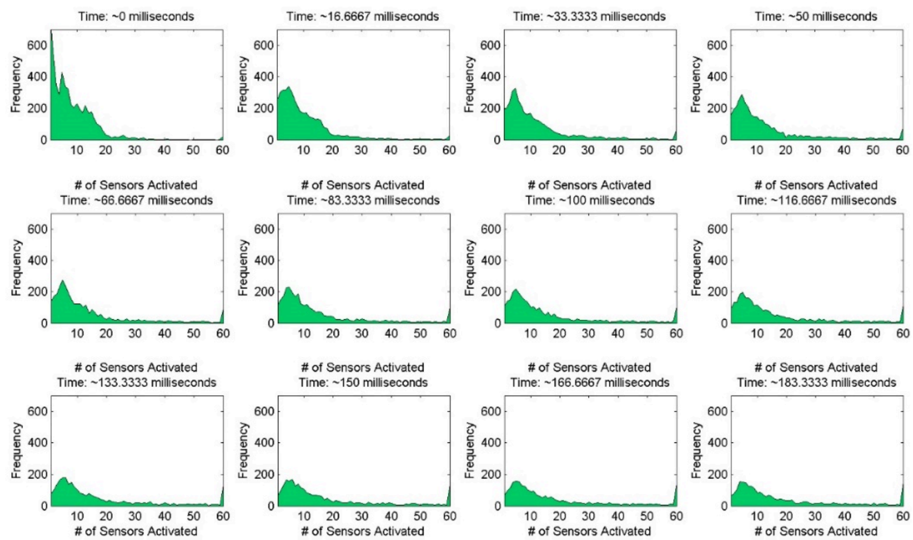


Figure 4.15. The area of each non-finger touch event on the first, second, third, etc. frame of its existence.

4.4.3.3 Hover-Based Rejection

The stylus height obtained from the motion capture system revealed that accuracy increases as the detectable hover height increases (Figure 4.16). At the hover height specified for Windows machines today, i.e., 20 millimeters, correct rejection occurs 88% of the time when intended touch is not generated (i.e., writing/tracing tasks) and 82% of the time when it is generated (i.e., annotation task). Whenever the detectable hover height is less than 20 millimeters, accuracy drops to 31-47% when there is intentional touch interaction, but only 48-87% when there is only unintended interaction (i.e., during writing and tracing). The disparity between the two scenarios is likely a by-product of the natural height at which the user holds the stylus while interacting with their non-dominant hand (see below). With advances in hover height detection, the accuracy of hover-based rejection would continue to increase, achieving approximately 96% accuracy at 100 millimeters and 100% accuracy at 200 millimeters with no intended interaction, and 82% accuracy at 100 millimeters and 83% accuracy at 200 millimeters when intended interaction is permitted. This is similar to what was found for the 'Reject All Touch Input' approach examined earlier.

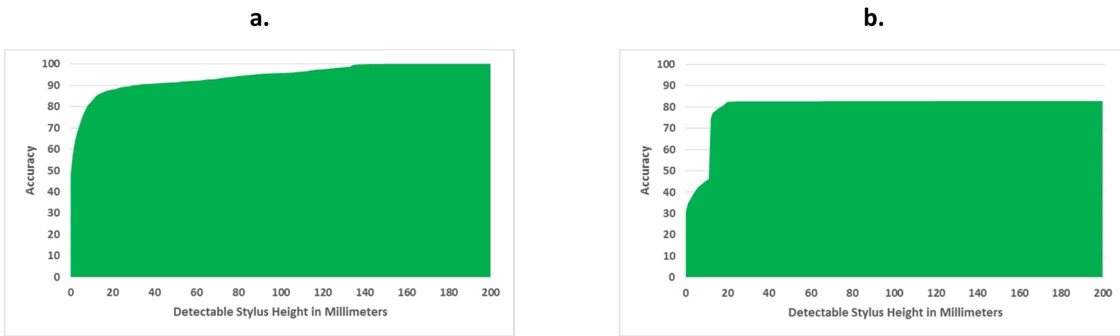


Figure 4.16. Accuracy results for the (a) Writing/Tracing tasks and (b) Annotation tasks while the current height of the stylus was used for rejection.

As it is likely that hover heights will increase beyond the 20 millimeter standard possible today, behavioral data gathered from the motion capture system suggests that hover will interrupt intentional uni-manual and bi-manual touch interaction. Across all of the gestures participants performed, the median nib heights were below 20 millimeters except for the 'Zoom Out' and 'Zoom In' gestures (27 millimeters and 30 millimeters, respectively; Figure 4.17). These nib heights largely reflect the range of stylus-manipulation techniques participants used. Most participants tucked the stylus behind their index finger to allow the index finger and thumb to gesture (80%). Such gestures were often restricted to low, short flicking finger motions that ensured that the stylus nib stayed close to the screen. Other participants grasped or clenched all of their fingers around the stylus (10%) or used their non-dominant hand to gesture (10%). The grasping motion allowed longer, more elongated gestures, which elevated the nib height, whereas when the non-dominant was used to gesture, the stylus and hand did not change position. Although gesturing occurred infrequently during inking, a majority of these gestures would be rejected, aggravating users. Given these natural behaviors, it is important to consider the impact that raising hover heights above their current threshold levels would have on interaction. As the detectable hover heights increase, designers need to carefully consider how they inhibit or support various bi-manual interactions.

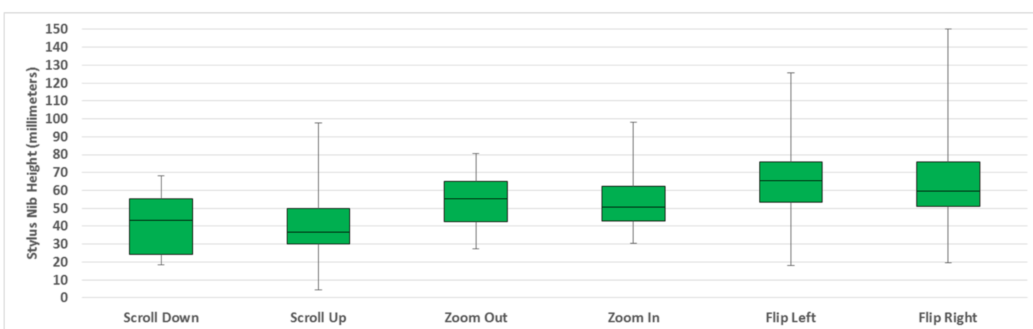


Figure 4.17. The median nib heights measured during the different gestures performed throughout the Annotation tasks. Error bars represent the standard error of the mean. The measurement error of the motion capture system was within ±1 millimeters.

4.4.3.4 Hand Occlusion Model

The Hand Occlusion Model, unfortunately, did not exhibit promising results (Figure 4.18). Across all detectable stylus heights, accuracy was slightly better during the annotation task. The hand posture parameters specified for the general model resulted in 18% accuracy (at 20 millimeters) when no intentional interaction was generated and 30% accuracy when intentional interaction was permitted (i.e., during the annotation tasks). If stylus detection was possible at higher heights, i.e., up to 200 millimeters, accuracy would marginally increase to 20% during writing and tracing, but stay at 30% while participants annotated. The disparity between the intentional and unintentional input scenarios (i.e., writing and tracing versus annotation) are likely due to the nature of the Hand Occlusion Model itself, which specifies the largest possible 'safe' area for intended interaction. Although the model allows for increased intended interaction (as evidenced by Figure 4.18b), its lack of support for unintentional interaction resulted in the poor accuracy performance that was found.

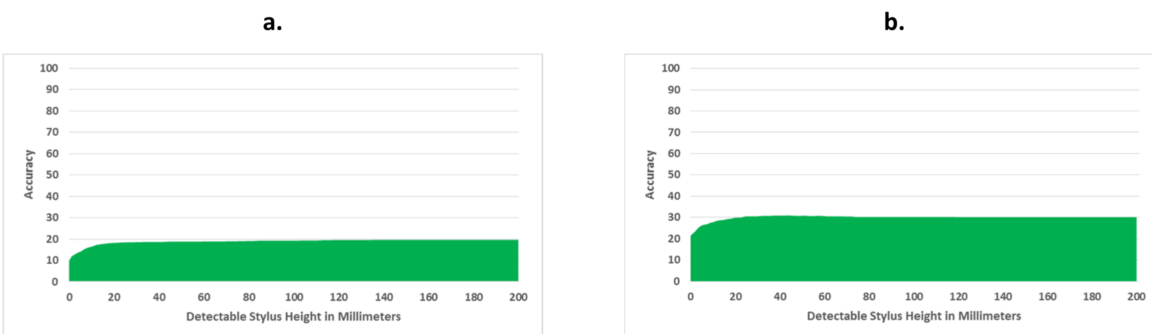


Figure 4.18. Accuracy results while using the Hand Occlusion Model during rejection for the (a) Writing/Tracing tasks and (b) Annotation tasks.

The overall inadequacy of the Hand Occlusion Model is likely based on two factors, the frequency of data available regarding the stylus location and the 'shape' of the activated touch input. An analysis of the motion capture data determined that a majority of the generated touch events were rejected due to insufficient stylus data. As per the requirements of the model, whenever touch input was initiated and the stylus was not present, the touch input was accepted (i.e., treated as a false positive). The motion capture data revealed that participants held the stylus nib approximately 22 millimeters (SEM = 9.27) above the surface when their palm first touched the screen.

When transitioning the stylus from one line to the next, the stylus was held higher, with the nib approximately 47 millimeters from the screen (SEM = 4.32). Such behaviors ensured that the stylus was outside the range currently detectable today and explain why more touch events were incorrectly accepted from 0-47 millimeters. When we consulted the video data, we also found that many participants used the bottom or side device bezel to stabilize the arm (Figure 4.19). This often resulted in the forearm crossing the bezel and generating touch input on the screen, much in advance of the wrist, palm, fingers, and stylus touching the screen. As the stylus height during this behavior far exceeded the detectable regions we tested, these touch events were incorrectly accepted, even when the stylus was detected at 200 millimeters.

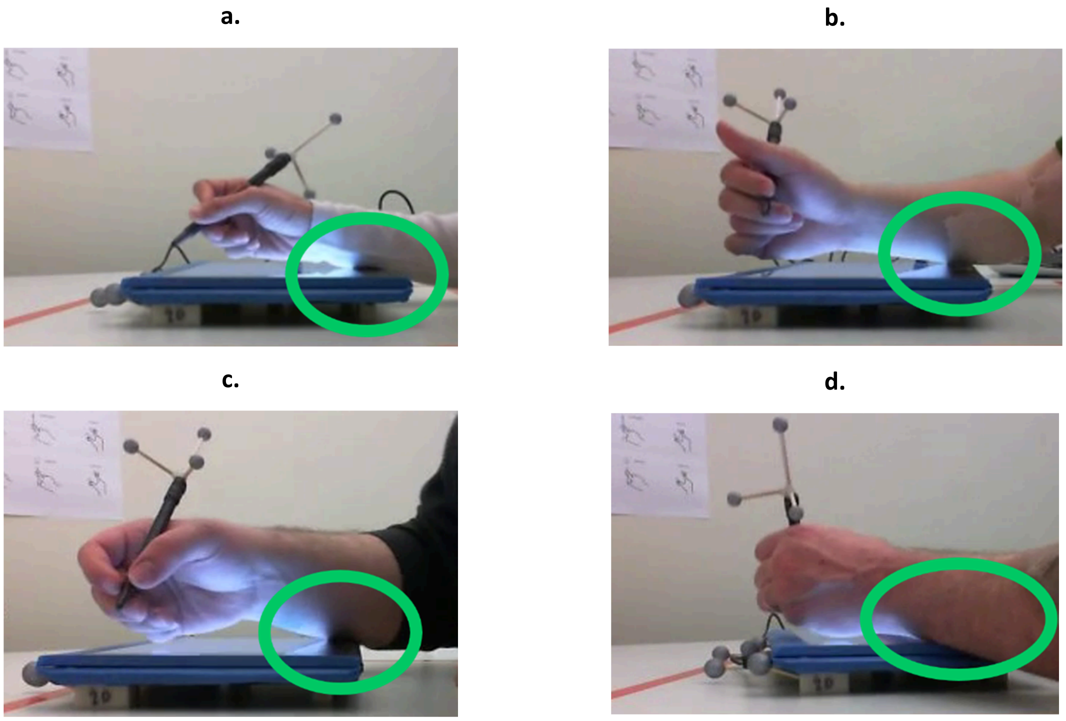


Figure 4.19. Examples of participants resting their forearm on the bottom bezel (a, b, c) and bottom of the touch screen (d) that generating touch input.

The generalized Hand Occlusion Model was also insufficient when one considers the shape of activation generated by participant's hand postures. Some participants used an inverted or 'hooked' hand posture (Levy & Reid, 1978; Figure 4.20a, c, d), while others used a more natural posture (Annett et al. 2014b), holding the stylus such that the nib was in a diagonal line with the palm and forearm (Figure 4.20b). These behaviors resulted in a majority of unintended touch input falling outside the generalized circle of the palm or rectangle of the forearm. While creating a personalized version of the hand model that could rotate farther from the nib (e.g., Figure 4.20c), translate closer to the nib (e.g., Figure 4.20b), or rotate the 'forearm' rectangle (e.g., Figure 4.20a, d) could eliminate some of these issues, the model would need to be continually updated on a stroke-by-stroke (Fitzmaurice et al., 1999), task-by-task (Annett et al. 2014b), or device-by-device basis. This would degrade the user experience and likely introduce additional application-specific latency into the latency pipeline.

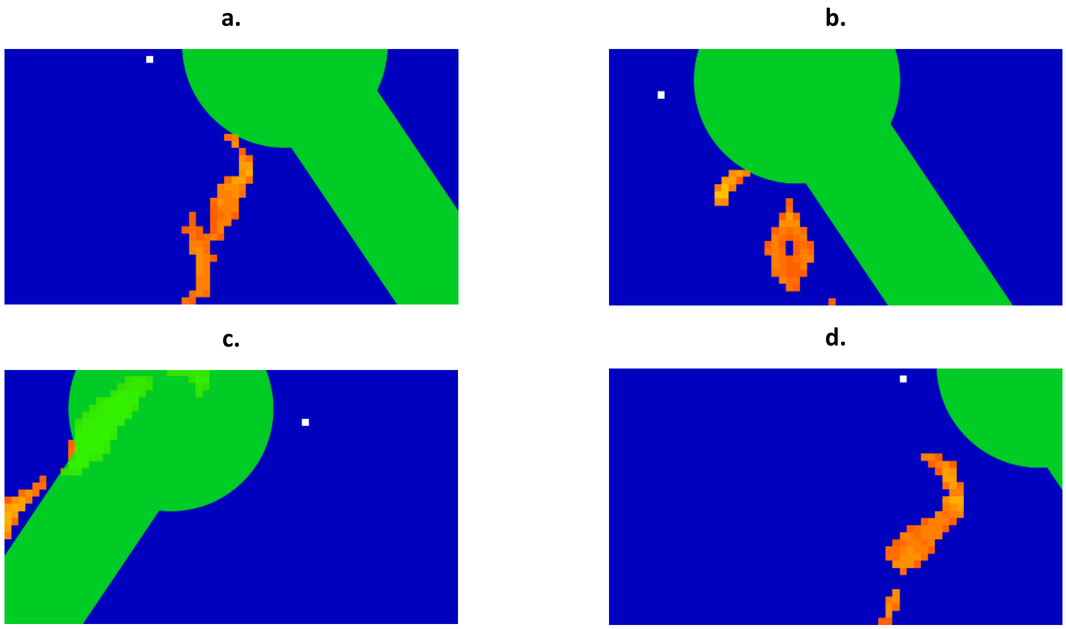


Figure 4.20. Examples of various hand postures that fell outside the generalized Hand Occlusion Model. The skin input is denoted in shades of orange, the stylus location is denoted by a white square, and the Hand Occlusion Model's abstracted circle and rectangle are green.

4.4.3.5 Static Rejection Regions

The static rejection region approaches performed quite well (Figure 4.21), partitioning the screen vertically (maximum rejection of 89% for writing/tracing and 78% for annotation) was superior to horizontally (maximum rejection of 69% for writing/tracing and 75% for annotation) and horizontally was superior to using a vertical-horizontal intersection (maximum rejection of 58% for writing/tracing and 44% for annotation). When considering the proportion of the screen 'safe' for interaction during writing/tracing, those covering two-thirds of the screen (maximum rejection of 89%) were superior to those covering half the screen (maximum rejection of 70%) and a third of the screen (maximum 49%). For the annotation tasks, those regions that covered two-thirds of the screen (maximum rejection of 78%) were superior to those covering half the screen (maximum rejection of 72%) and a third of the screen (maximum 66%). Across all manipulations, a maximum of 87% correct rejection was found for the writing/tracing tasks and 79% for the annotation tasks. There thus appears to be a general trend towards screen partitions that rejected touch input from the largest vertical area. Vertical rejection regions likely work well because they capitalize on the natural ergonomics and kinetics of the lower arm. When on the screen, the wrist, forearm, and hand generally fall along a diagonal or vertical line, and are normally oriented vertical instead of horizontal.

a. **b.**

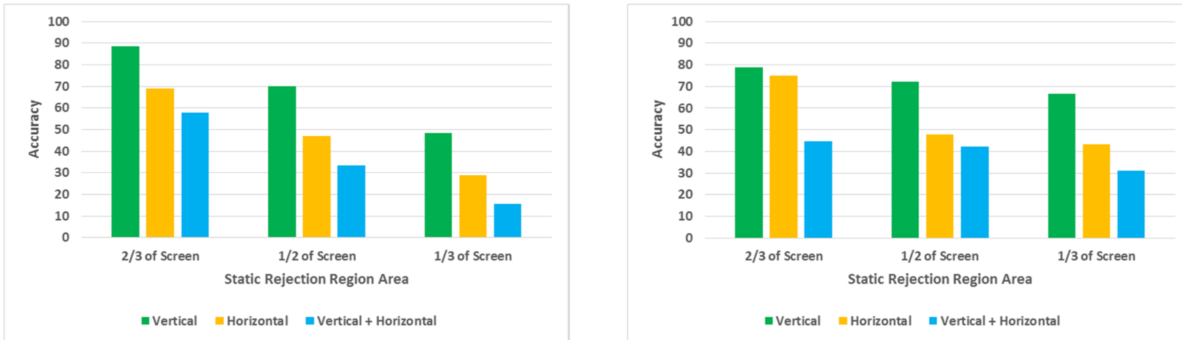


Figure 4.21. The accuracy found for various static region-based approaches while participants were (a) Writing/Tracing and (b) Annotating.

To better understand the performance of the various configurations, the locations of all touch inputs generated by participants were graphed using heat maps (Figure 4.22). In the writing tasks, most touch events occurred in the bottom right quadrant of the screen, where the forearm, wrist, and palm rested while right-handed participants were writing (Figure 4.22a). Surprisingly, touch input was not simply mirrored to the bottom-left for the left handed participants. Instead, left-handed participants exhibited unique behaviors, generating many more touch events along the left and top edges. In the tracing tasks, touch input for both handedness populations was much more dispersed and found in many different locations of the screen (Figure 4.22b). The hot spots found in the center and along the left and bottom right bezels are indicative of areas where users rested their arms on the screen while completing the drawing tasks in each quadrant. As much less interaction was required in the annotation tasks, it is unsurprising that input was found in the same areas for all participants (and somewhat mirrored for left-handed participants, Figure 4.22c, d). As interaction was restricted to key areas for each participant, touch input patterns were concentrated in these locations.

These task-dependent differences help explain why static rejection regions are a popular and reasonable first choice for unintended touch. They harness the task context and hypothetical behavior found with these tasks to eliminate the majority of unintended touch events, e.g., those occurring in the bottom right for right-handed participants. As writing and tracing required interaction in all areas of the screen, it is unsurprising that the larger rejection regions worked better than the smaller ones. Although the typical distance between the hands, while performing bi-manual actions, has yet to be explored, the similar results found between the writing/tracing and annotation accuracy results suggest that the 2/3 vertical split left enough room for users to interact bi-manually when they needed to (i.e., during the annotation task).

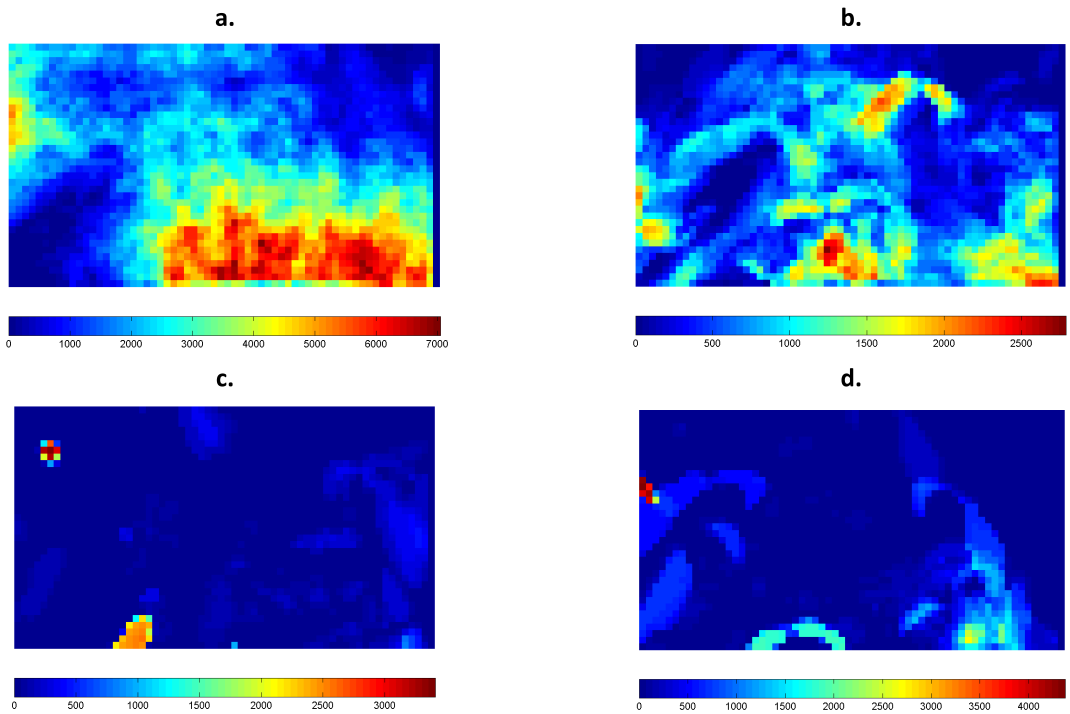


Figure 4.22. Heat maps of the touch input generated by all participants for the (a) Writing task, (b) Tracing task, (c) first Annotation task, and (d) second Annotation task. The heat map data has not been normalized.

4.4.3.6 Stylus-Based Rejection Regions

The use of the stylus to specify the location of various rejection regions showed a similar trend to the static rejection regions: the use of a vertical region was better than a horizontal region, and both were better than a horizontal-vertical intersection (Figure 4.23) Based on the distribution of touch events presented in Figure 4.22, it makes sense that approaches that were most accommodating to unintended touch and the natural behavior or users would be successful. Across both writing/tracing and annotation, increasing the detectable stylus height (i.e., the instances where touch events had corresponding stylus information), lead to a slight increase in performance. Given the technical limitations of stylus detection today, i.e., up to 20 millimeters, a maximum of 62% accuracy was possible with the vertical partitioning for the writing/tracing tasks and 46% for the annotation task. If a tablet could detect the stylus up to 200 millimeters from the screen, accuracy increased to 68% and 50% respectively.

What was found to differ with stylus-based regions is that accuracy decreased when intentional touch was supported (i.e., annotation tasks). The nature of the touch and stylus input was quite different during the annotation tasks. Instead of being able to drag the hand around the screen from left to right, top to bottom, the annotation task naturally encouraged replanting behavior (Annett et al., 2014b), i.e., participants placed their hand on the screen, performed an action, picked it up and held it in the air, replanted it in a different area of the screen, and so on. This intermittent behavior resulted in many more touch events that could not be classified

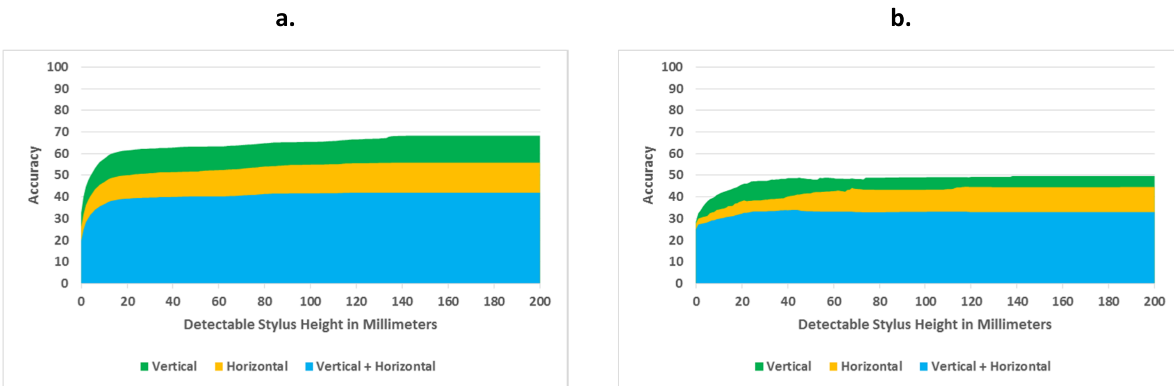


Figure 4.23. The accuracy of stylus-based rejection while using different rejection regions and when the stylus is detectable from different heights. Image (a) contains data from the Writing / Tracing tasks and (b) contains data gathered during the Annotation tasks.

because information about the stylus was missing. Similar to the Hand Occlusion Model, the success of stylus-based rejection regions depends on the availability of the stylus location. Although the location of the stylus was tracked at all times, the higher the stylus is in the air, the more uncertainty there is regarding where it would eventually land on the screen due to the range of motion possible via the degrees of freedom afforded by the elbow, wrist, and fingers. This uncertainty manifested itself in decreased accuracy.

The decreased accuracy across all tasks exhibited by the stylus-based rejection regions (compared to the hover and static rejection regions), is also due to the variability of hand postures that could be used while interacting. Although information about the stylus location could provide context as to where the user is interacting, the video and touch data revealed that such context information is not very useful. An analysis of the hand and wrist postures determined that the use of natural and inverted hand postures placed contacts outside the specified rejection area (Figure 4.24). On many occasions, the stylus location fell in line (Figure 4.24b), far below (Figure 4.24a, d, e, f), or to the side (Figure 4.24c) of the touch input in question. For this reason, the largest rejection region, the vertical screen partition, was the most tolerant to non-traditional hand postures and rotations of the wrist. Ignoring only the touch events in the bottom or bottom right corner was too inflexible, as the hand reorients as it moves towards the bezels and bottom of the screen. The use of larger 'safe' areas accommodates the potential movement of the wrist and hand throughout the task.

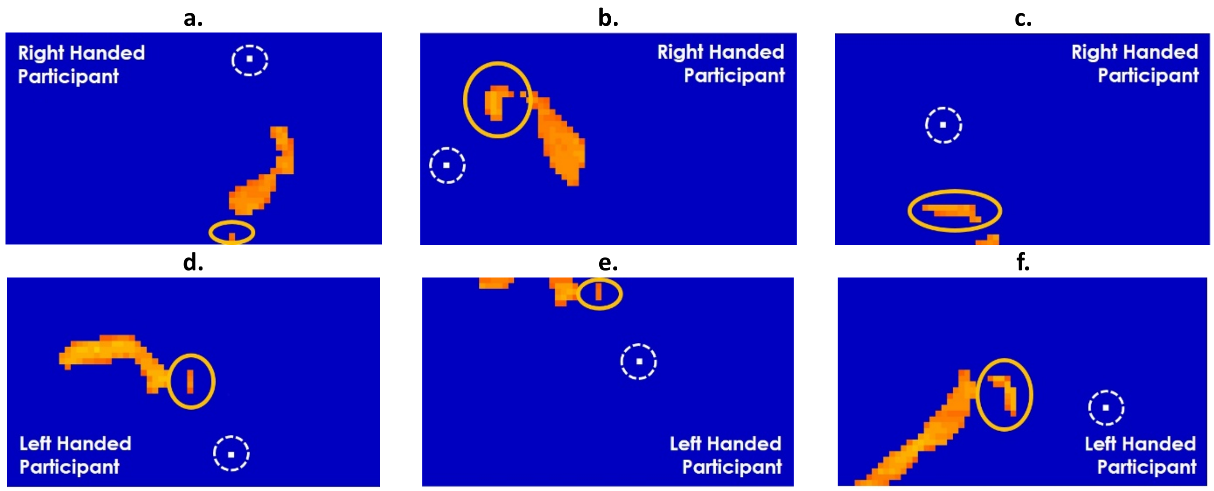


Figure 4.24. Examples of various touch inputs generated by participants that would be incorrectly accepted when the stylus location is used to define the rejection region. The skin contact is denoted in shades of orange, the stylus location is denoted and circled by a dashed white line, and the touch input under consideration is circled with a solid orange line.

4.4.3.7 Stylus-Based Rejection Regions Using a Buffer

The results obtained with the stylus-based rejections using a buffer demonstrate the same general trend as the static rejection regions and stylus-based rejection regions: the use of a vertical partioning worked best for rejection (i.e., maximum accuracy of 80% accuracy while writing/tracing, and 71% accuracy while annotating; Figure 4.25). Given the distribution of touch events in Figure 4.22 and the vertical region being attuned to the anatomy of the hand, these results underscore the importance of harnessing natural user behavior and observations for unintended touch.

The results also illustrate that as the size of the buffer increased, so did the accuracy (e.g., the vertical partition saw an increase from 62% to 72% Figure 4.25a; from 65% to 72% Figure 4.25b; from 68% to 80% Figure 4.25c; from 64% to 71% Figure 4.25d). When we consider this 'best case', i.e., vertical partition, and the stylus is detectable from 0 – 20 millimeters, a 50 millimeter buffer will result in 72% accuracy when only unintentional touch is considered (i.e.,, writing and tracing) and 71% accuracy when intended and unintended touch is allowed (i.e., annotation). If the stylus can be detected from 0 – 200 millimeters, accuracy increases slightly to 80% while writing/tracing and stays the same while annotating. Although we did not statistically analyze the writing/tracing and annotation tasks, there was little difference between them, save the slightly higher variance between the spatial partitions during annotation. The results thus demonstrate that the larger the buffer region, the greater the increase in accuracy, illustrating that the buffer succeeded in accommodating the variety of hand postures likely during inking and that rejection was largely not influenced by the desire to perform intended versus unintended touch.

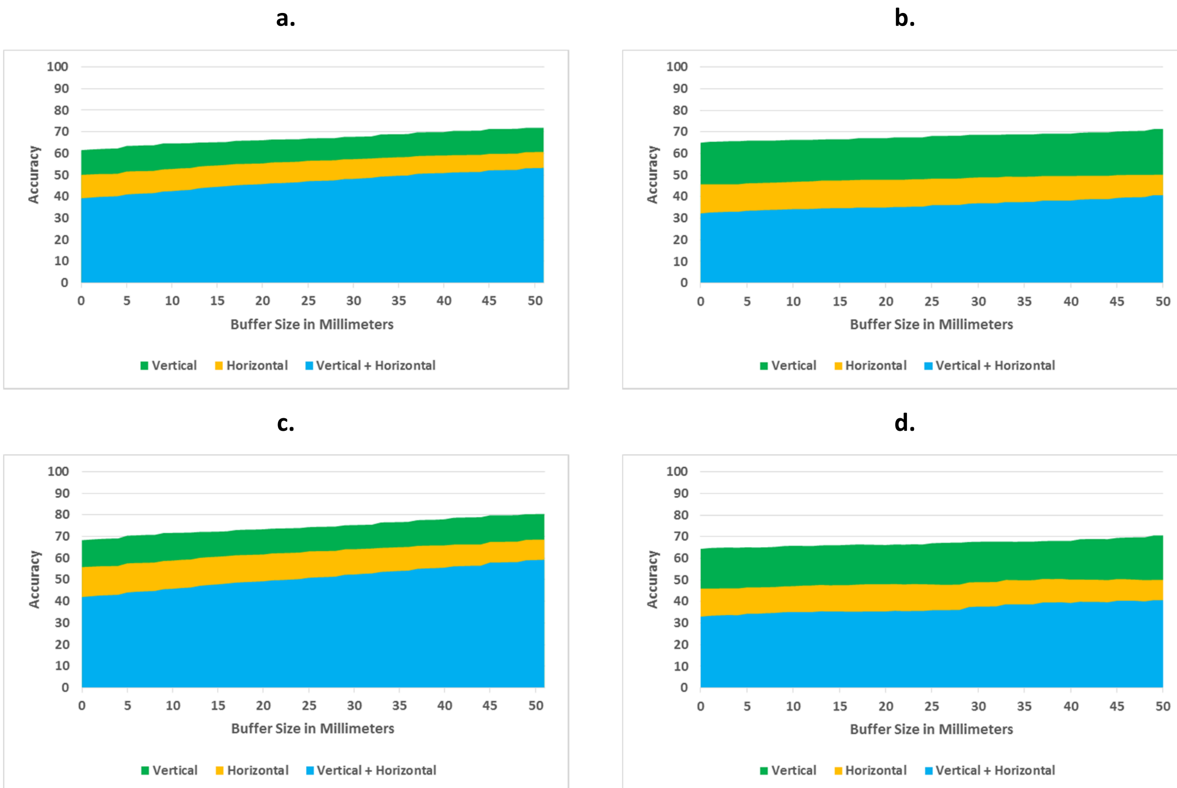


Figure 4.25. Accuracy of the stylus-based rejection region approaches with different rejection partitions while (a,c) Writing/Tracing and (b,d) Annotating. Images (a,b) illustrate the accuracy when the stylus is detectable from 0-20 millimeters and (c,d) illustrate the accuracy when the stylus is detectable from 0-200 millimeters.

The touch-based data from Figure 4.22 demonstrated that participants may not always use a “natural” hand posture. It is thus not surprising that an approach that accommodates variations in posture, grip, or general anatomy is more successful than one that cannot. When we reassess the instances where user behavior resulted in the stylus-based rejection regions failing (from Figure 4.25), the use of a buffer would eliminate the unintended touch created by the palm wrist, fingers, and forearms (Figure 4.26). While the buffer does increase the likelihood that hand posture can be accounted for, it does, of course, require that the stylus location is always known. Although the use of a buffer resulted in better performance, there were still a number of unintended touch events that were accepted due to the stylus being above or outside the detectable stylus range (and thus the rejection region location and boundaries could not be determined).

Given that the palm or forearm often touches the surface before the stylus, these results provide even more support for acquiring data as soon as possible. If the stylus location was always known, possibly using technology other than what is available today, this approach could work better because it can harnesses contextual information about where the user will be interacting. When combined with handedness, it becomes very powerful, allowing for a variety of hand postures and supporting many bi-manual input techniques.

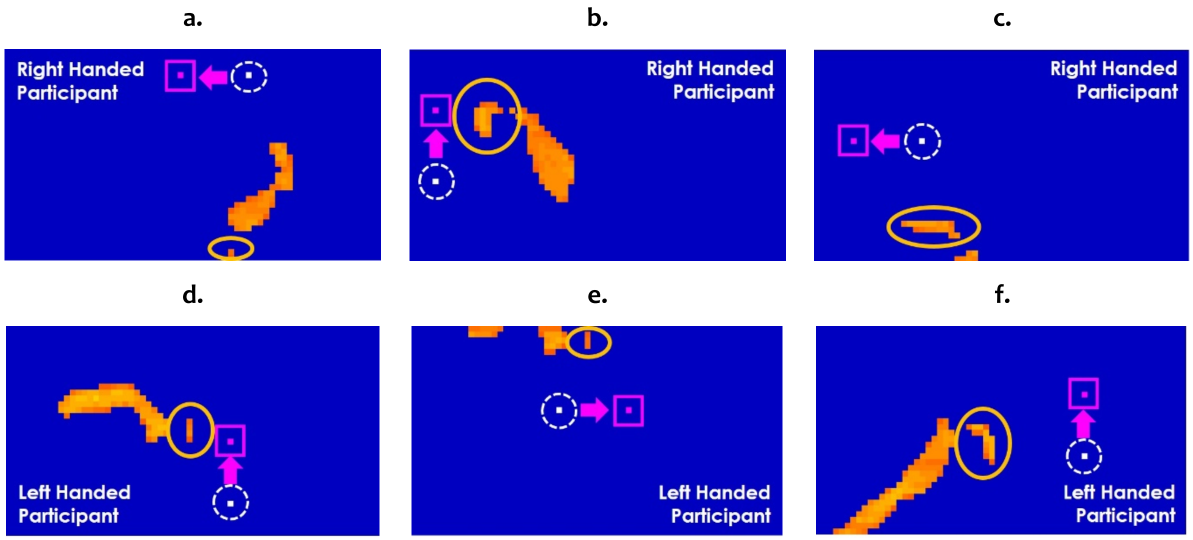


Figure 4.26. A demonstration of how the use of a buffer can accommodate many of the unintended touch events that were incorrectly accepted in Figure 4.24. The skin input is denoted in shades of orange, the current touch blob under consideration is circled using a solid orange line, the detected stylus location is denoted by a white square and circled via a white dashed line, and the resulting stylus location after a 50 millimeter buffer was added is displayed in purple and highlighted by a purple square.

4.4.3.8 Aggregated Contingency Table

As many parameters and variations were evaluated for each algorithm, it would be unwieldy to present the raw contingency tables for each variation that was examined. Thus, an aggregated summary of the raw contingency table values (i.e., number of the true positives, true negatives, false positives, and false negatives) can be found in Table 4.2. The presented data illustrate the best results found for each algorithm given the technological capabilities of today (i.e., the detectability of the stylus up to 20 mm) and the most reasonable parameters.

4.4.4 Comparative Evaluation of Algorithms

The exploration thus far has focused on algorithmic approaches that have many possible variations and options. Participants' behavior helped explain why some of the approaches were successful in natural inking settings, and found the same trends regardless of the task and underlying amount of intentional versus unintentional touch input. At the same time, the results underscored the importance of understanding user behavior during inking and the complexity of unintended touch. What is missing, however, is a comparison and understanding of the approaches that are best suited for unintended touch, equally weighing accuracy and the limitations of each approach, the required functionality, and how they would impact the cohesion of user experiences across the stylus and tablet ecosystem.

Table 4.2. A summary of the aggregated raw contingency tables found for each algorithm that was evaluated. The percentage of true positives (TP), true negatives (TN), false positive (FP), and false negatives (FN) are presented within the brackets.

Algorithm	Writing / Tracing				Annotation			
	TP	TN	FP	FN	TP	TN	FP	FN
No Algorithm	0	5085 (100%)	0	0	0	1005 (83%)	0	213 (17%)
Contact Area[1]	0	2596 (51%)	2489 (49%)	0	98 (8%)	362 (30%)	643 (53%)	115 (9%)
Hover[2]	0	4474 (88%)	611 (12%)	0	204 (17%)	800 (66%)	205 (17%)	9 (1%)
Hand Occlusion Model[3]	0	926 (18%)	4159 (82%)	0	212 (17%)	152 (12%)	853 (66%)	1 (0%)
Static Rejection Region[4]	0	4506 (89%)	579 (11%)	0	42 (3%)	916 (75%)	89 (7%)	171 (14%)
Stylus-Based Rejection Region[5]	0	3128 (62%)	1957 (38%)	0	212 (17%)	345 (28%)	659 (54%)	2 (0%)
Stylus-Based Rejection Region with Buffer[6]	0	3653 (72%)	1432 (28%)	0	208 (17%)	403 (33%)	602 (49%)	5 (0%)

[1] Contact area of four sensors.
[2] A detectable stylus height of 20 millimeters.
[3] A detectable stylus height of 20 millimeters.
[4] A vertical rejection region covering 2/3 of the screen.
[5] A detectable stylus height of 20 millimeters with a vertical rejection region.
[6] A detectable stylus height of 20 millimeters with a vertical rejection region and a 50 mm buffer.

4.4.4.1 Methodology

As each of the approaches use different types of data and rely on different assumptions, an analysis of how they fared against each other, when the technological restrictions of tablets and styli were removed, was performed. A comparison of each approach, from the perspective of today's technology and the technological improvements likely in the foreseeable future, was then performed using the optimal parameters and results for each algorithm from Section 4.4.3 (Table 4.3). Given that the same trends were found across the writing/tracing and annotation tasks, the data was aggregated to remove task as a factor. Note that even with technological advancements, not all approaches will improve, e.g., increases in sensor density will increase the number of sensors that are activated uniformly.

Using the parameters in Table 4.3, a two-way ANOVA with the factors of Functionality (present, future) and Algorithm (i.e., contact area, hover, hand model, static rejection region, stylus-based rejection region, and stylus-based rejection region with buffer) was run using the accuracy values determined from the contingency table for each participant.

Table 4.3. The parameters used in the comparison of the unintended touch algorithms given current and future functionality.

Algorithm	Present Parameters	Future Parameters
Contact Area	▪ 4 sensors	▪ 4 sensors
Hover	▪ 20 mm hover height	▪ 200 mm hover height
Hand Model	▪ Stylus location known within 20 mm of surface	▪ Stylus location known within 200 mm of surface
Static Rejection Region	▪ Vertical rejection region covering 2/3 of the screen	▪ Vertical rejection region covering 2/3 of the screen
Stylus-Based Rejection Region	▪ Vertical rejection region ▪ Stylus location known within 20 mm of surface	▪ Vertical rejection region ▪ Stylus location known within 200 mm of surface
Stylus-Based Rejection Region with Buffer	▪ Vertical rejection region ▪ Stylus location known within 20 mm of surface ▪ Buffer of 50 mm	▪ Vertical rejection region ▪ Stylus location known within 200 mm of surface ▪ Buffer of 50 mm

4.4.4.2 Results

The ANOVA results demonstrated that Functionality significantly influenced accuracy ($F_{1,15} = 5.0$, $p < .05$) as did Algorithm ($F_{2.3,35.0} = 30.6$, $p < .001$; Figure 4.27). There was no evidence of a Functionality by Algorithm interaction ($F_{1.4,20.8} = 2.6$, $p = 0.112$). Increases in technological functionality and capabilities led to small, but significant, enhancements in the performance of some algorithms that were evaluated (*Present*: M = 62.73, SEM = 3.00; *Future*: M = 65.56, SEM = 3.20; $p < .05$), namely the Hover (i.e., 6% improvement in accuracy), Stylus-Based Rejection Region (i.e., 5% increase in accuracy), and Stylus-Based Rejection Region with Buffer (i.e., 4% increase in accuracy) approaches.

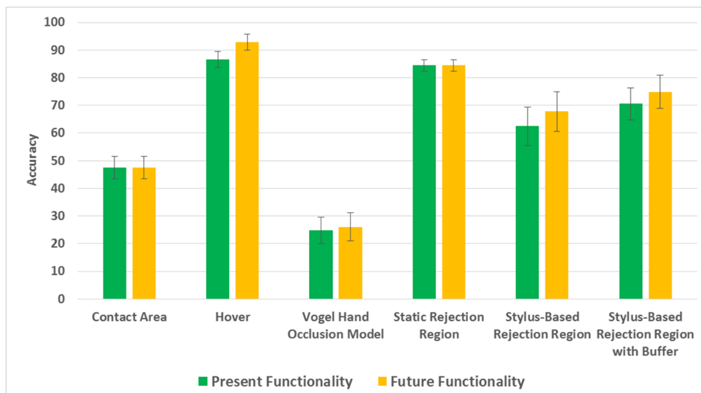


Figure 4.27. A comparison of the six unintended touch algorithms that were evaluated using current and future functionality.

Post-hoc Bonferroni-adjusted paired t-tests demonstrate that some approaches are much better suited for unintended touch than others (Table 4.4). Approaches that made use of stylus data or supported larger 'safe' areas of the screen (i.e., Hover, Static Rejection Region, Stylus-Based Rejection Region, and Stylus-Based Rejection Region with Buffer) outperformed those using generalized data sources with smaller 'safe' areas, such as Contact Area or the Hand Occlusion Model. It thus appears that the size of the 'safe' versus 'unsafe' areas and the use of contextual and peripheral information to specify the location of these areas (e.g., current task, stylus height, stylus location) are important to unintended touch.

Table 4.4. The results of the pairwise comparisons between various unintended touch algorithms (* denotes significance).

Pairwise Comparison	ΔM	SEM	*p*
Contact Area vs Hover	-42.3	5.5	0.015 *
Contact Area vs Hand Occlusion Model	21.9	5.3	0.015 *
Contact Area vs Static Rejection Region	-37.1	4.6	0.015 *
Contact Area vs Stylus-Based Rejection Region	-17.7	7.9	0.63
Contact Area vs Stylus-Based Rejection Region with Buffer	-25.3	6.8	0.03 *
Hover vs Hand Occlusion Model	64.3	6.5	0.015 *
Hover vs Static Rejection Region	5.3	2.3	0.57
Hover vs Stylus-Based Rejection Region	24.7	7.6	0.075
Hover vs Stylus-Based Rejection Region with Buffer	17.1	5.8	0.15
Hand Occlusion Model vs Static Rejection Region	-59.0	6.3	0.015 *
Hand Occlusion Model vs Stylus-Based Rejection Region	-39.6	6.6	0.015 *
Hand Occlusion Model vs Stylus-Based Rejection Region with Buffer	-47.2	6.4	0.015 *
Static Rejection Region vs Stylus-Based Rejection Region	19.4	8.2	0.48
Static Rejection Region vs Stylus-Based Rejection Region with Buffer	11.8	6.6	1.44
Stylus Rejection Region vs Stylus-Based Rejection Region with Buffer	-7.6	2.3	0.09

4.4.4.3 Discussion

The comparative analysis found a significant difference between contact area and the hand occlusion model. Given that the hand occlusion model relies on the stylus location and has a very narrow spatial area within which touch input is allowed, it is not surprising that the hand occlusion model performed worst of all the approaches. As shown in the individual analysis, participants' hand postures were quite variable, rarely conforming to the abstract generalization that the hand occlusion model required. Although one could make use of a trained, adaptive hand model, the task-by-task heat maps show that, depending on the requirements of a task, participants exhibit drastically different hand patterns from start to finish, especially in the more free-form tasks, such as tracing and annotating. Apart from the problem of variable hand behavior, personalized hand models would need to be relearned when transitioning to a new device, form factor, or task. This would likely involve a calibration phase, something many developers are very resistant towards.

By comparison to the other four approaches, the use of contact size performed no better than chance, due in large part to the length of time necessary for each touch event to grow to a distinguishable size before rejection could occur. For the contact area approach to be more successful, touch input would need to be provided before skin touches the screen, i.e., when the hand or forearm is within the 'hover' state. This would then allow the disambiguation and rejection processes to begin before touch input activates the digitizer and possibly soon enough for correct rejection and acceptance to occur.

As there are no significant differences between the top four approaches, it cannot be claimed that one is superior to the others. There are a number of marginally significant results, however, that can be used to better understand where to devote research resources in the future. The hover-based approach is significantly, or marginally significantly, better at unintended touch than the other approaches. In essence, the use of hover creates the largest rejection region: the whole area of the screen. This makes it the most tolerant to non-traditional hand movements and positions but limits its usefulness for scenarios where bi-manual interaction is desired. As the rejection region gets smaller, either because it is bounded by the stylus location, the buffer dimensions, or the screen partitioning, the success of space-based approaches decreases. The results found with the stylus-based and static rejection regions support this conclusion.

The marginal difference between the buffer and non-buffer stylus-based rejection regions demonstrates that allowing a 'spatial cushion' around the stylus is a simple, but effective way to accommodate rotations of the wrist and palm. Furthermore, the relatively small rejection area allows for some bi-manual interaction, which is not supported by the hover-only approach. Future approaches that use stylus location should consider integrating a similar spatial cushion. Based on the heat map distributions presented earlier, the static rejection region approach did a good job of rejecting unintended touch in tasks where the hand is anchored on the screen in a predictable manner, i.e., while writing. As stylus information is not needed, rejection occurs quickly, and does not result in false positives whenever the stylus location is not present. With applications such as note-taking on capacitive devices, it would be beneficial to use such a region, as it would prevent most unintended touch interaction. Coupling a visual representation with the rejection region allows users to understand where they can rest their palm without disrupting their work and increases the likelihood they write as naturally as possible.

4.5 Overall Discussion

The present exploration indicates that there are many elements that contribute to unintended touch, making the problem, and choice of an unintended touch solution, very complex. Although our underlying data set does not contain an equal number of intentional versus unintentional touch events, we believe that the data set and results are representative of the distribution of such events common during real world inking tasks, wherein the frequency of unintended touch almost always outweighs intended touch. For example, while taking notes during a lecture, it is likely users would only intentionally touch the screen whenever they needed to scroll the page. Such behavior results in a drastically smaller proportion of intentional compared to unintended touch events, due in large part to the frequency of the hand dragging or being replanted on the screen (Annett et al., 2014c) compared to page scrolling. As such, approaches that focus on, or are slightly biased toward identifying unintended touch instead of intended touch (e.g., hover, static rejection region, and so on) would perform better (e.g., stylus-based rejection region, stylus-based rejection region with a buffer, and so on).

The similar trends found in the individual algorithm analysis, exemplifies that the more opportunities and support are provided for intentional touch, the more contextual knowledge regarding user behavior is required. If one considers a sketching application that allows users to zoom in and out via a pinching gesture to add details or perform shading, for example, successful approaches to eliminate unintended touch will need to be increasingly cognizant of the multitude of hand postures possible, the frequency of screen reorientations (Fitzmaurice et al., 1999), the likely diverse distribution of unintended touch events around the screen, and the increased frequency of intentional touch events depending on the level of detail required by the user. Although the present study did not evaluate such scenarios, and we chose to eliminate touch-based feedback for the user, it is easy to see the importance of further exploring and understanding natural user behavior to inform the design of unintended touch algorithms. The abundance of pen and touch work performed over the last few years has generated many novel interaction techniques and possible gesture taxonomies, but unfortunately the most basic of knowledge is missing, e.g., about the relative spacing between the hands while interacting bi-manually, the orientation of the hands and stylus relative to each other and the screen, the prevalence of bi-manual over uni-manual interaction, and so on. Such knowledge will greatly improve the models of inking behavior that can be build and integrated within unintended touch solutions.

In what follows, the accuracy results and observed participant behavior are situated within the context of real world demands and technological implications. Potential avenues for future research are also explored, focusing first on the benefits of different data streams and then detailing various hardware advancements and software modifications that could greatly improve the problem of unintended touch.

4.5.1 Factors Important for Unintended Touch

The goal of any pen and touch system or application is to allow a transfer of the seamless, interleaved bimanual interaction found in the physical world to the digital world. It is thus imperative that developers be mindful of the limitations that a given algorithm may place on interaction. In addition to accuracy, there are several other factors that need to be considered when understanding the approaches that could be viable in the future (Table 4.5). In some contexts, such as a writing application where a user may only need to scroll periodically, a hover-based approach may work well, whereas in a sketching application, interleaved zooming and inking may be required frequently, resulting in that same hover-based algorithm to underperform and cause frustration for uses.

Algorithms implemented at the application level also allow context-specific rejections and application-specific interactions, but do so at the cost of a fragmented user experience that is inconsistent across applications and tasks. Firmware and operating system promote a unified user experience, but must be primitive, making use of a more generic, context, and functionality-free designs. Depending on the needs of a developer or the end user population, the most accurate approach may be too restrictive for the ecosystem.

Table 4.5. A comparison of unintended touch approaches along various dimensions that influence interaction.

Approach	Implementation Level	Active Stylus Required?	Handedness Information Required?	Support for Various Hand Postures?	Supports Bimanual Interaction?	Display Orientation and Rotation Required?
Contact Area	Firmware or Operating System	No	No	Yes	Yes	No
Hover	Operating System	Yes	No	Yes	No	No
Hand Occlusion Model	Operating System or Application	Yes	Yes	No	Yes	No
Generic space-based (e.g., rejection regions)	Firmware	No	Yes	No	No	Yes
Stylus space-based (e.g., rejection region with stylus and / or buffer)	Operating System or Application	Yes	Yes	No	Depends on buffer size	Yes

Approaches that make use of handedness or harness data made available by an active stylus lead to enhanced accuracy but fragment the user experience. They are inappropriate for capacitive touchscreen devices that only support passive styli. If a lower price point is targeted, the cost of integrating an active stylus may outweigh the benefits perfect rejection would provide to users. The desire to capitalize on and integrate spatial data adds an additional requirement that a device must be aware of orientation, rotation, and handedness to be successful. As the detection of rotation and handedness are still in the research stage, one may be forced to sacrifice the expectations of certain populations or use cases.

Until unintended touch can be solved completely, it is imperative that the algorithms and metaphors employed are discoverable and easy to understand. When the data collection experiment began, many participants were surprised that it was technically possible to rest the skin on the screen while interacting, "it allowed me to interact so naturally – it was like I was writing on paper!" In the post-experiment debriefing, participants who had prior experience with inadequate unintended touch reaffirmed the results found by Annett et al. (2014b), expressing great frustration with current devices today: "I never know what I do that causes the screen to zoom or random marks to appear", "It is so difficult to write with an iPad because some applications work fine but in others my palm messes everything up and I get really upset" and "Your system worked great! My hand never made a mark … I wish others were like that". Users were very quick to determine if an approach does not work (i.e., they can't place their hand), but are likely to have difficulties determining why or how to adapt their behavior to overcome unintended touch. Given the fragmentation and variety of unintended touch algorithms used today, these comments highlight the need for consistent approaches to unintended touch. Designers and developers must explore a variety of use cases when designing solutions and mindfully consider ways to either alert the user if an approach falters or provide fluid, intuitive ways to overcome and avoid errors.

Given these factors and the accuracy results, it becomes clear that there is no easy solution for unintended touch. When evaluating and implementing potential solutions to unintended touch, designers and developers must prioritize and balance the requirements of their end users, the interaction they desire, and the consistency of their applications or operating systems within the tablet ecosystem and across their application suites.

4.5.2 Use of Additional Data Streams

Although the approaches tested here used only hover and stylus location, there are many other approaches and streams of data that could be harnessed for unintended touch in the future. While the use of the hover state alone for rejection is not advocated, the 'pre-touch' information it provides is valuable. If decisions can begin before the user touches the screen, with either the hand or the stylus, the likelihood of incorrectly rejecting or accepting touch will decrease. Recently, some mobile phones have begun to ship with support for 'touch hover', enabling for in-air gesturing with the finger or hand (Sony, n.d.; Samsung, n.d.; STMicrocontrollers, n.d.). Such data could greatly improve unintended touch, as information could be obtained about the type of contact before it reaches the screen. This information could boost the performance of many of the algorithms evaluated here, in essence allowing for two-stage disambiguation and classification. Given the widespread availability of such information in the mobile world, it is only a matter of time before such information will be available on larger form factors such as tablets.

There are other stylus-based streams of data that could also be useful for unintended touch. Tablets today supporting active styli provide 'pre-touch' information in the form of an x/y hover cursor. While they have information regarding the height of the stylus from the screen, this information is not yet available for use by developers. The availability of this information would provide a better picture of the time remaining before the stylus, hand, or forearm touch the screen. This information would also allow for a plethora of proximity-based interaction techniques. Integration of additional sensor information, such as pen roll or tilt, or the pressure or grip on the barrel of the pen (Hinckley et al., 2013, Song et al., 2011) could also be useful when combined with the application context. Developers need, however, to be mindful when integrating such information, as the use of additional sensors could come at an increased cost, processing time, latency, or more complex signal processing, depending upon when such information was available and were it was integrated.

When considering devices that only have a capacitive input layer, the increased popularity of auxiliary passive styli such as Pencil, Ink, Adnoit Jot Pro, and Pogo Connect, may become an important key to solving unintended touch on such devices (Annett, 2014a). If these styli were also able to provide information about the stylus location, similar to the Wacom Inkling (n.d.) or active styli, it may be possible to bring a coherent, consistent user experience to both active and passive systems.

4.5.3 Necessary Hardware- and Software-Based Improvements

As alluded to, the availability and frequency with which information becomes available for unintended touch will greatly influence the approaches that are possible. Today, there is a very limited time frame to collect information and make a decision regarding touch input. As system latencies decrease it may be possible to allocate a few newly 'freed' frames to unintended touch, delaying the rejection decision for each touch event by a few frames. This would allow touch input sizes to stabilize, more stylus data to become available, and temporal-based approaches to be implemented. It will, however, be a balancing act between maintaining a latency-free experience prone to false touch events versus inducing latency and eliminating unintended touch.

Modifications of firmware or operating system APIs could allow developers to receive a 'rejection confidence level' along with each touch input or event. Such information could be based on a variety of factors including sensor magnitudes, duration of activity, etc. and would not introduce additional latency. There are also possibilities for systems without disambiguation as well. Visualizations could indicate the confidence a system or application has for various touch events it thinks may be unintended (e.g., those along the borders of an application or at spatially disjoint locations). Systems could vary the opacity of strokes, use colored overlays, or display specific cursors to illustrate how confident they are in the intentionality of strokes without removing them from the screen. This could help make the user aware of behaviors that cause stray markings or unintended navigations and possibly help them avoid such behaviors in the future. Visualizations could also indicate which touch points are likely from the pen and which are from the finger. Pairing these visualizations with a brushing, sweeping, or swiping gesture could help users to quickly erase these extraneous markings.

The present study required that the tablet was flat on the table and oriented towards the user in landscape mode, with only a small rotation permitted. Although rotation was corrected for in the data set, tablets today do not have such capabilities. If tablets could report their current rotation and orientation instead of simply reporting 'portrait' or 'landscape', developers could harness this information to improve the robustness of their unintended touch solutions whilst opening the possibility for new interaction techniques and functionalities. There are many instances where the tablet may not be flat (e.g., when held, sitting vertically, or resting on the lap (Wagner et al., 2012)), when the user could be interacting upside down (e.g., by reaching across the table to interact with someone else's device), or the tablet could be continually rotated, similar to the behavior found with artists while sketching (Fitzmaurice et al., 1999). In these scenarios, algorithms making use of spatial information would fail, even the ones that rely on hand models. The present study did not focus on gathering information on the behavioral and postural changes that tablet rotation and orientation would produce. For this reason, it is imperative that developers carefully consider how such behaviors may influence unintended touch in these contexts.

The study also assumed that the tablet was handedness-aware and ignored the details necessary to collect this information. While some applications utilize user interface widgets and dialogs to gather such information, there has already been much work on the automatic detection of handedness (Dang et al., 2009; Ewerling et al., 2012; Ramakers et al., 2012; Wang et al., 2009; Zhang et al., 2012). In addition to these approaches, the data collected from the digitizer suggests that it may be possible to detect handedness at the firmware level, by simply using the movement of touch blobs over time or their spatial relation to the stylus location. A firmware-based approach would eliminate the need for external cameras, augmented setups, sensors, or user interface interventions while providing valuable information for unintended touch, occlusion-free interfaces, and novel adaptive interface widgets. This exploration is left for future work.

As detailed in Section 4.2.2, companies including Synaptics (Coldewey, 2011; Sage, 2011), Atmel (n.d.), and Samsung (SmartKeitai, 2013) have announced improvements in digitizer signal-to-noise ratios that have enabled better disambiguation between thin, narrow inputs (e.g., from a stylus) and larger, wider input (e.g., from fingers and palms). Once such approaches and technology exits the prototype stage and differentiated passive styli become available, it would be useful to explore more dynamic approaches to unintended touch, especially in light of the increased vocabulary of gestural commands that are becoming possible with these devices (Marquardt et al., 2011; Wigdor et al., 2011; Yan et al., 2013).

4.6 Summary

With stylus-enabled devices today, the inking experience is largely marred by the problem of unintended touch. Digitizers and applications have a difficult time distinguishing those touch events that are desired from those that are unintended. When unintended touch persists, it hampers a user's ability to interact naturally with their device and restricts the transfer of their learned behaviors from paper and pen to the digital world. This work has presented an in-depth analysis of unintended touch from both an algorithmic and a user behavior perspective.

The use of a motion capture system, along with raw data from touch and stylus digitizers allowed for an evaluation of a variety of approaches to unintended touch. In the evaluation, many novel behaviors were observed: the maximum time possible for unintended touch decisions, the typical height of the stylus during uni-manual or bi-manual gestural interaction, and the pattern and location of touch events. These observations were found to be important for understanding the effectiveness of various approaches and can also be useful for pen computing in general.

The analysis revealed that some current approaches, that limit the 'safe' areas of the screen, such as the use of hand models or the size of touch contacts do not fare well, while others that allow for larger rejection regions such as hover or those specified by the location of the stylus, demonstrate much better performance. Deriving a general solution to solve unintended touch is difficult and influenced by factors relating to desired interaction and functionality, and information availability. Rejecting unintended touch should be made much easier in the future, however, when pre-touch, handedness, device rotation, and orientation information are available for integration. Such data will allow ample time to integrate and synthesize data before the skin has touched the surface of the screen, thus enhancing the pen and touch user experience.

Chapter 5

Stylus Accuracy

The diversity within the stylus and tablet ecosystem as demonstrated in Chapter 2 has resulted in markedly different user experiences, especially with regards to accuracy. It is quite common for there to be an incongruence between the location where a display draws ink and the location at which a user expects their ink to appear. While such inaccuracy has been acknowledged for many years (Annett et al., 2014a; Badam et al., 2014; Read, 2006), there are still many open questions regarding: i) the sources of errors, ii) the relationships and relative importance of these sources, and iii) methods to prevent such errors during selection and inking tasks. Occlusion (Lee et al., 2012; Ramos et al., 2007; Vogel & Balakrishnan, 2010a) and visual parallax (Annett et al., 2014a; Badam et al., 2014; Ramos et al., 2007; Vogel & Balakrishnan, 2010a) have long been regarded as the only factors influencing accuracy, resulting in little attention devoted to understanding or identifying if any other features are implicated in the inaccuracy experienced today. The present work provides the first in-depth analysis of the behavioral, technological and stylus-based factors influencing accuracy, identifying and classifying eight other factors that influence (perceived) accuracy.

As it is infeasible to tease apart the relations between all the factors at once, two experiments assessed a subset of these factors (i.e., visual feedback, natural hand postures, and nib aesthetics). By systematically evaluating factors using standardized tasks and measures, a richer understanding of accuracy was attained, extending past traditional factors that have been the focus as of late. The quantitative results and participant comments underscored the importance of visual feedback and nib aesthetics while devaluing the role that natural hand postures play in accuracy.

5.1 Sources of Inaccuracy

Although many hardware and software advances have been made in the decades since digital styli were introduced, users still cite inaccuracy as a major source of frustration. In the research literature, many novel software-based widgets have been proposed to overcome inaccuracy or enhance precision while pointing and selecting objects (Blanch et al., 2004; Grossman & Balakrishnan, 2005; Ramos et al., 2007). Within commercial settings, most efforts focus on hardware advancements involving calibration and visual parallax (Vogel & Balakrishnan, 2010a). While such advancements and techniques have improved accuracy, such improvements are often incremental, have yet to achieve mainstream adoption, or cannot be applied to scenarios outside of selection, such as writing and sketching.

The factors that influence latency were identified and collected from the literature, keeping both industry, research, and end-user interests in perspective (Figure 5.1). The results of this exploration identified 10 factors that influence accuracy: *visual parallax, friction and surface texture, sensor linearity and configuration, calibration techniques, transducer parallax, the design of the stylus, occlusion, hand postures, visual feedback,* and *user experience and expectations*. These factors have been loosely organized into three categories: *underlying technology, stylus design,* and *user interaction*. The identification and awareness of these factors should stimulate the research and industry communities to broaden their foci and increase their efforts to understand and address these issues.

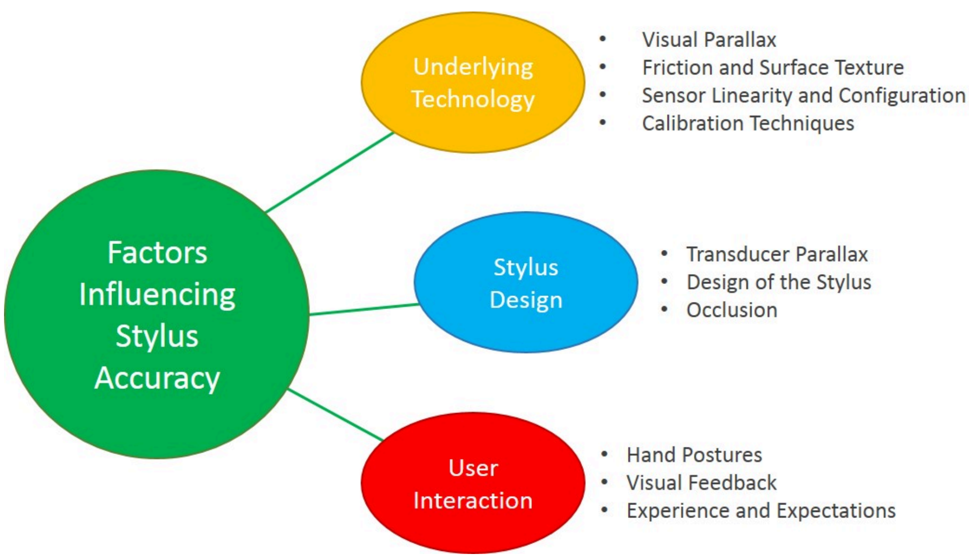


Figure 5.1. The three categories of factors likely implicated in the poor accuracy experience had by users today.

5.1.1 Underlying Technology

The choices made by manufacturers, in addition to current technological limitations, have resulted in at least five different elements that influence accuracy. One of the foremost factors is *display or visual parallax* (i.e., the distance between the surface of the screen and the underlying sensing digitizer (Lee et al., 2012; Ramos et al., 2007; Vogel & Balakrishnan, 2010a)). The thicker the glass used in the display and the farther the digitizer is from the surface, the more inaccuracy. While advancements had been made to combine touch and stylus digitizers and to bond digitizers directly to the screen to reduce this distance, it still remains difficult to determine the precise location of the stylus.

Friction and surface texture have also been cited as possibly important to the accuracy perceived and errors experienced (Chapter 2; Annett et al. 2014a; Mohr et al., 2010; Sun et al., 2012). With most device manufacturers utilizing glass for the surface of the display, it becomes easy to lose control of the stylus. While the use of glass materials may be aesthetically pleasing, when coupled with a plastic nib, immense fine motor control and increased effort are needed by the user to ensure that the stylus lands on, and moves across the screen as desired.

For consistent calibration and accuracy, *sensor linearity* and *configurations* must be designed to ensure sufficient signal redundancy in all areas of the screen (Ward & Philips, 1987). Sensors must extend past the edge of the display to allow for the accuracy in the corners to match accuracy in the center of the display. This introduces additional costs and engineering efforts. Additionally, when using an active stylus with electromagnetic sensing, the introduction of novel magnetic fields from a metal table or tablet case can cause drift, making the device appear miscalibrated and inaccurate, seemingly at random.

Accuracy can also be influenced by the *calibration techniques* that align the sensor and display spaces (Ward & Philips, 1987). If an inadequate or improper calibration is performed by the user, in the factory, or in the firmware, the displayed location of the stylus will appear offset from its real location. While users may adapt to offsets, switching to a new stylus or device, or any new recalibration that is performed will require a relearning of the current offset, which may not be linear.

5.1.2 Stylus Design

As users are very particular about the writing and sketching implements they prefer, it should be expected that the *design and aesthetics of a stylus* would influence precision and accuracy as well. Although many have evaluated user comfort and preferences for various barrel designs, lengths, and weights (Goonetilleke, Hoffman, & Luximon, 2009; Park, Kim, & del Pobil, 2011; Wu & Lu, 2006) little attention has been given to other equally important facets of stylus design. As demonstrated in Chapter 2, the material requirements of the nib are also likely to influence accuracy (Chapter 2; Annett et al. 2014a). When nib materials are malleable, such as those found with passive styli, the deformation caused by pressing the stylus to the screen creates occlusion and ambiguity with regard to the sensed location of the pen. Similar to touchscreen interaction (Holz & Baudisch, 2011), confusion about the sensed location of the stylus will result (i.e., center of the nib, towards the top, bottom right, etc.). Additional factors such as the taper and design of the barrel (Ward & Philips, 1987), the degree to which the nib protrudes from the stylus, and the shape of the nib are also likely to be implicated in inaccuracy (Chapter 2; Annett et al. 2014b).

Occlusion is additionally often cited as a major source of inaccuracy (Chapter 2; Annett et al. 2014a; Badam et al, 2014; Ramos et al., 2007; Vogel and Balakrishnan, 2010a). Similar to traditional pens, the presence of the nib itself can occlude content. The diameter disparity between traditional pens and pencils (i.e., approximately 0.7 mm) and digital styli (i.e., approximately 1.6 millimeters to 7 millimeters), creates much ambiguity regarding the sensed position. Techniques such as callout widgets (Vogel & Balakrishnan, 2010a) have proposed spatially offsetting the location of on-screen content. The PhantomPen (Lee et al. 2012) has proposed virtually reconstructing the stylus nib to allow hidden content to become visible.

An additional form of parallax, *transducer parallax*, can also influence accuracy. With the use of electromagnetic or active capacitance to detect an active stylus, the circuitry required for sensing is contained within the stylus itself. As transducers themselves are large, they are relegated up into the barrel of the stylus. This creates an offset between the tip of the nib and stylus sensor itself (Ward & Philips, 1987). When the stylus is held vertically, inaccuracy is not apparent, but when the stylus is held at an angle (as is common), the sensor-tip offset is manifested onscreen.

5.1.3 User Interaction

Accuracy is rarely hypothesized to be influenced by user behaviors or user factors. As Chapter 2 illustrated, *hand posture* adaptations are unique to digital devices and were the by-product of inadequate unintended touch and inadequate surface textures (Annett et al. 2014a). These unnatural postures were the cause of the inaccurate and messy content because they did not provide the wrist and hand stability needed while writing and sketching. Matulic and Norrie (2012) made a similar observation while users performed stylus-based tasks on a multi-touch tabletop and prevented from resting the wrist, forearm, and elbow. Both of these studies identified the important role that support can play with accuracy, but failed to investigate further.

Visual feedback can also play a role in accuracy, as it has long been found to be beneficial for aiming movements (Carlton, 1981). Tablets that have the ability to detect the position of stylus before it touches the screen can provide visual feedback in the form of an onscreen cursor. Although such cursors do not exist with traditional pen and paper, exploiting the hover state provides users the opportunity to correct their movements before any permanent actions are taken (Buxton, 1990). It is unknown to what degree users depend on this information or how much the user experience degrades when such feedback is unavailable.

In addition to hand posture, prior *experience and expectations* with pen and tablet systems can influence perceived accuracy as well. Many users erroneously assume styli are sensed by a tablet using resistive, pressure sensitive means, similar to early touch screens. Those with passive styli experience are aware they can be imprecise. Others may have had poor experiences with prior systems and automatically adapt their behavior to overcome issues that no longer exist. Each of these scenarios decreases user confidence with the system, possibly affecting unconscious motor commands as well. This lack of confidence and the frequent comparisons users make between pen and paper and poor digital experiences could manifest in the content created (Chapter 2; Annett et al., 2014a).

5.2 Experiment 1: Visual Feedback and Hand Posture

From a pragmatic perspective, evaluating all the aforementioned factors that could influence accuracy is infeasible. Thus, the first experiment selected a subset of unexplored factors to assess, namely visual feedback and hand posture. Although the hover cursor is a visual feedback mechanism to overcome inaccuracy, it is unknown how much the user experience degrades when the feedback the hover cursor provides is unavailable. Does the hover cursor slow users down because they wait to see it? As the hover cursor is not available with pen and paper, is it distracting? By modifying the presence and absence of the hover cursor, an understanding of its importance to accuracy, speed, and user preferences could be attained.

As Chapter 2 identified, the behavioral adaptations to overcome unintended touch and inadequate surface textures are common and frequent complaints of users (Annett et al., 2014a; Matulic & Norrie, 2012). They were thus the second factor of interest. When resting the hand on the screen, one is able to isolate the movement from the wrist upwards to the shoulder, decreasing the noise inherent in the motor movements needed to control the stylus (Bernstein, 1967). When the wrist or hand is elevated above the surface, users must make compensatory movements to stabilize forearm, upper arm, and stylus. These additional movements may add noise (i.e., wobbly lines and messy content) to the system. By explicitly allowing or disallowing participants to rest their forearms and wrists, the ‘noise’ introduced by such behavioral adaptations could be measured.

5.2.1 Participants

Sixteen participants (6 female) were recruited to participate in the study (*M* = 25 years, *SD* = 4 years, range = 18-34 years). All participants were right handed. Participants were naive to the purpose of the experiment and the majority had prior experience using touchscreen tablets. Three participants used a stylus on a regular basis. Participants received a $20 honorarium for participating in the 30 minute experiment.

5.2.2 Experimental Apparatus

A Samsung Series 7 Business Slate, running Windows 8.1 with a resolution of 1366 x 768 pixels, was placed in front of participants in landscape orientation (Figure 5.2). Participants could reorient and reposition the tablet as needed to ensure that they were comfortable, but were required to keep it on the desk. Custom C# and WPF controlled the presentation of each task and recorded the data generated by participants.

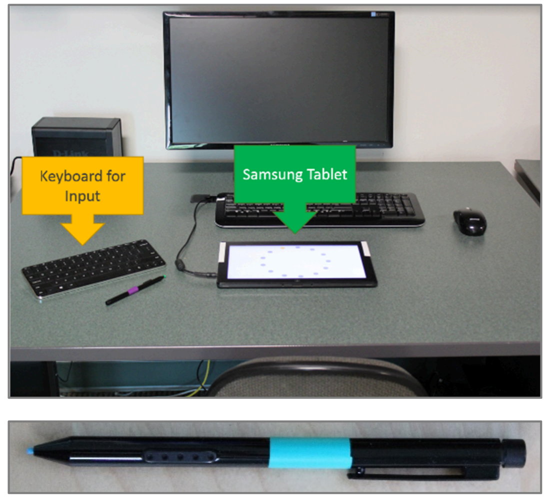


Figure 5.2. (Top) The experimental setup with the Samsung Slate and keyboard used in the writing task. (Bottom) The Surface Pro pen used in the experiment.

For inking and selecting targets, participants used a Surface Pro stylus (Figure 5.2). The stylus had blue nib that was 1.6 millimeters thick and protruded 1.8 millimeters from the stylus. Throughout the experiment, all touch input was disabled to ensure participants could touch the screen without fear of consequence. Additionally, all stylus button input was disabled and the traditional Windows 8 diamond-shaped hover cursor was replaced with a visually similar cursor that could be controlled programmatically.

5.2.3 Experiment Conditions

Two factors were tested in the study. The first factor, Hover Cursor, manipulated the feedback available whenever the stylus was in the hover state. In the hover-visible condition, the hover cursor was visible whenever the stylus was less than 20 millimeters above the screen and disappeared when the stylus was in contact with the screen; in the hover-invisible condition, the cursor was never visible.

The second factor, Hand Resting, manipulated participants' ability to rest their wrist, fingers, knuckles, etc. on the screen. In the touch condition, participants were instructed to interact normally with the device, resting their skin if desired. In the no-touch condition, were explicitly instructed to avoid touching the screen but could rest their elbow on the table. Across both conditions, participants were encouraged to take breaks to reduce fatigue.

5.2.4 Tasks and Procedure

A majority of the stylus-based activities can be distilled into one of two types of interaction: selection or stroking (inking). As such, participants completed two tasks from the ISO 9241-9 (2002) and ISO 9241-411 (2012) standards.

In the first task, *selection*, participants performed the ISO 9241-411 2D multi-directional tapping task (2012). This task was chosen because it simulated the selection of targets or icons in an interface. Participants were presented with a radial array of targets and were required to select the highlighted target as quickly and accuracy as possible (Figure 5.3). Three target diameters (i.e., 12, 32, and 53 pixels) and three target amplitudes, or inter-target distances, (i.e., 140, 463, and 683 pixels) were used, resulting in 9 experimental conditions. As each condition contained 13 targets, participants made 12 selections before the next trial began. The index of difficulties ranged from 1.86 to 5.86 bits (ISO/TS 9241-9:2000E, 2012).

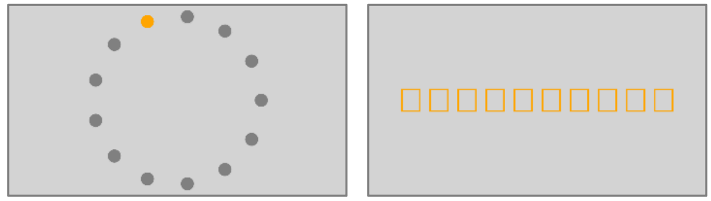


Figure 5.3. Screenshots of the selection (left) and writing (right) tasks performed by participants.

In the *writing* task, participants performed a variant of the ISO 9241-9 free-hand input test (2002). Nine rectangles appeared on the screen and participants were asked to write the digits from 0 to 9, one number per rectangle (Figure 5.3). Participants were instructed to write each digit such that it was centered and completely contained within each rectangle. After participants wrote the last digit, they were instructed to press the space bar on a Microsoft Wedge Mobile keyboard to advance to the next trial. Three rectangle widths (i.e., 32, 52, and 78 pixels) and three inter-rectangle distances (i.e., 1, 36, and 53 pixels) were used, resulting in 9 experimental conditions.

Participants were seated in a computer chair in front of a standard office desk throughout the experiment. At the beginning of each task, participants were given a short demonstration of the task to be performed. As multiple tasks and conditions were evaluated, task presentation order and the order of trials was counterbalanced using a Latin square design. At the conclusion of the experiment, participants answered a short questionnaire on noticeability, accuracy, and preferences.

5.2.5 Measures and Data Analysis

Although many methods can be used to measure the accuracy of target selections, few are also applicable to written content (Read, 2006). Two measures of accuracy were thus chosen, one coarse and one fine. The coarse measure was an adaptation of one of Mackenzie et al.'s (2001) pointing accuracy measures. For the selection task, *error* was defined as the number of attempts necessary to tap each target. For the writing task, each participant could write each letter on a single attempt so the number of strokes that fell either fully or partially outside the target rectangles was used.

The fine accuracy measure, *movement offset*, was an adaptation of Mackenzie et al.'s (2001) movement offset measure. Movement offset for the selection task was the deviation in millimeters from the center of the target circle to the user's selection location. For the computation of movement offset in the writing task, the bounding box surrounding each digit was determined and the distance in millimeters from the center of the box to the center of the target rectangle was computed.

In pilot studies, it was observed that many users modified the *pressure exerted* with the stylus when asked to hold their hand in the air. Based on these observations, the pressure exerted also was analyzed. For both tasks, the pressure recorded by the stylus was in the range of 0 - 255.

Lastly, *duration* assessed the impact of hand posture and visual feedback on input speed. For the selection task, the average time to complete each of the 12 selections was computed. For the writing task, duration was measured from the beginning of the first stroke to the end of the last stroke.

5.2.6 Results

A within-subjects, repeated measures ANOVA with Hover Cursor (levels: Present, Absent) and Hand Resting (levels: Allowed, Not Allowed) was performed. As the computations for each measure varied between tasks, (e.g., duration averaged across the 12 selections for selection but total duration for writing), each task was analyzed separately. For the questionnaire data, responses were encoded on a 7-point Likert scale, with 1 corresponding to "Strongly Disagree", 4 to "Neutral", and 7 to "Strongly Agree". All responses were compared to the neutral response using one-sample T-tests.

5.2.6.1 Errors

With Hover Cursor, participants made fewer targeting attempts when the cursor was present compared to absent (Figure 5.4a; $F_{1,15} = 23.7$, $p < .001$, $\eta^2 = .38$). Hand Resting was not found to influence targeting accuracy ($F_{1,15} = .3$, $p = .62$, $\eta^2 = .00$) and no interaction was found between Hover Cursor and Hand Resting ($F_{1,15} = 1.9$, $p = .19$). The writing task demonstrated similar results: fewer errors were found for the Hover Cursor factor when it was present (Figure 5.4b; $F_{1,15} = 9.8$, $p < .001$, $\eta^2 = .18$). Hand Resting was not found to influence error ($F_{1,15} = 0.7$, $p = .71$, $\eta^2 = .00$) and no interaction was found between Hover Cursor and Hand Resting ($F_{1,15} = 0.4$, $p = .84$). Thus, visual feedback appears to decrease error and helps targeting movements and inking. While the accuracy of the initial targeting movement should not be related to one's ability to keep their strokes within the target rectangle, the results suggest otherwise.

5.2.6.2 Movement Offset

In the Hover Cursor conditions, participants' selections were more accurate when the cursor was present than absent (Figure 5.4c; $F_{1,15} = 15.7$, $p < .001$, $\eta^2 = .49$). With Hand Resting, participants exhibited less variability when their hand was in the air (Figure 5.4c; $F_{1,15} = 5.2$, $p < .05$, $\eta^2 = .03$). No interaction was found between the Hover Cursor and Hand Resting factors ($F_{1,15} = 0.9$, $p = .36$). While it appears that both visual feedback and hand posture are important, the effect sizes suggest visual feedback is more important.

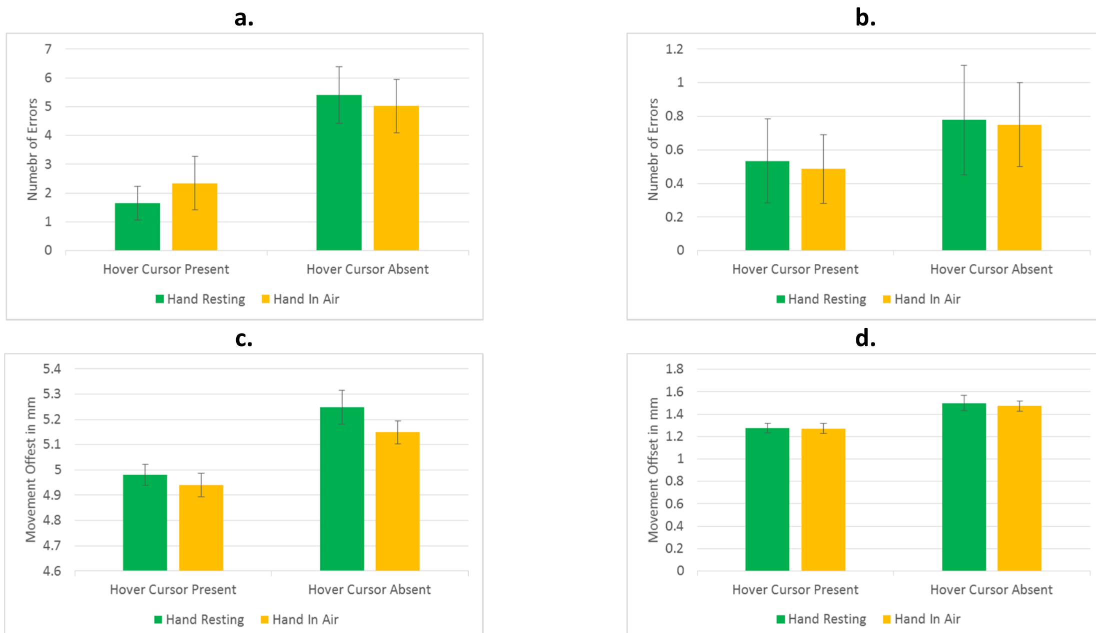


Figure 5.4. (a) The error found for the selection task, (b) the error found for the writing task, (c) the movement offset for the selection task, and (d) the movement offset for the writing task. The error bars represent the standard error of the mean.

While writing, the Hover Cursor factor resulted in digits that were significantly more centered when feedback was present (Figure 5.4d; $F_{1,15} = 44.2$, $p < .001$, $\eta^2 = .50$). Hand Resting had no influence on accuracy, ($F_{1,15} = 0.15$, $p = 0.70$, $\eta^2 = .00$) and no Hover Cursor and Hand Resting interaction was found ($F_{1,15} = 0.3$, $p = .62$). Echoing the selection results, the stability afforded by the screen is of less importance to accuracy than the visual feedback provided by the cursor.

5.2.6.3 Pressure Exerted

For the selection task, Hand Resting was found to influence the pressure exerted. Participants exerted more pressure when their hand was on the screen than in the air (Table 5.1; $F_{1,15} = 8.9$, $p < .01$, $\eta^2 = .23$). The Hover Cursor condition did not influence the pressure exerted ($F_{1,15} = 2.0$, $p = .18$, $\eta^2 = .03$), nor was there a Hover Cursor by Hand Resting interaction ($F_{1,15} = 1.0$, $p = .33$). Hand Resting did influence the pressure exerted while writing, with significantly more pressure being exerted while the hand rested on the screen (Table 5.1; $F_{1,15}$ = 8.6, p < .01, η^2 = .29). The Hover Cursor factor did not influence the pressure exerted ($F_{1,15}$ = 3.1, p = .10) and a Hover Cursor by Hand Resting interaction was not found ($F_{1,15}$ = 1.0, p = .33). These results complement the other accuracy measures and suggest that the use of pressure to assess perceived accuracy may be useful.

Table 5.1. The results for the Duration and Pressure Exerted measures for Experiment 1.

		Duration in Milliseconds				Pressure Exerted			
		Selection		Writing		Selection		Writing	
		M	SEM	M	SEM	M	SEM	M	SEM
Hand Resting	Cursor Present	900	36	6660	312	27.73	4.32	126.23	8.39
	Cursor Absent	1020	53	6481	332	29.48	3.34	126.54	9.25
Hand In Air	Cursor Present	1006	47	7649	371	19.77	3.42	99.13	7.93
	Cursor Absent	994	40	6945	299	23.75	3.15	107.42	7.06

5.2.6.4 Duration

During the selection task, the Hover Cursor factor influenced speed. Participants completed their selection faster when the hover cursor was present than when it was absent (Table 5.1; $F_{1,15}$ = 4.3, p < .05, η^2 = .07). Hand Resting did not influence task time ($F_{1,15}$ = 1.6, p = .24, η^2 = .03) and no interaction was found between the Hover Cursor and Hand Resting factors ($F_{1,15}$ = 7.0, p = .06).

While writing, the Hover Cursor and Hand Resting factors influenced speed. Participants were faster when allowed to rest their hand on the screen than when they were not allowed to do so (Table 5.1; $F_{1,15}$ = 11.7, p < .01, η^2 = .27). Participants were also slower when the hover cursor was present (Table 5.1; $F_{1,15}$ = 13.5, p < .01, η^2 = .10). No interaction was found between the Hand Resting and Hover Cursor factors ($F_{1,15}$ = 4.2, p = .06). It thus appears that the feedback from the hover cursor slowed users down, possibly because they waited to receive feedback before interacting. When it comes to inking speed, the effect sizes suggest that hand posture is more important than feedback.

5.2.7 Discussion

The results demonstrated that hand posture plays surprisingly little role in accuracy. When the wrist is in the air, the increased degrees of freedom (e.g., fingers, wrist, and forearms) do not introduce significant instability or inaccuracy, contrary to hypotheses from prior work. The resting of the elbow (and forearm) on the table provided enough stability for participants to perform the necessary movements. Resting the hand on the screen, while useful for writing, was not helpful while selecting targets. This is likely because the larger inter-target distances placed targets outside the natural range of motion of the wrist, hence participants had to lift their hand to complete the task anyways.

Hand resting led to participants to exert 22% more pressure while targeting and 11% more pressure while writing. Although resting the hand should have led to decreased pressure, the differences may be due to the slipperiness of the plastic nib on the glass surface. When the hand is not available to stabilize the stylus, the loss of control possibly leads to a loss of exerted force. Participants were also found to demonstrate an 11% increase in speed while writing with their hand on the screen. As holding the hand in the air is uncomfortable and unnatural, participants were likely much more careful to ensure they would not have to repeat a movement.

Many participants felt it was distracting ($M = 5.07$, $SD = 1.83$; $t(15) = 2.26$, $p < .05$) to hold their arm in the air and commented on the fatigue and discomfort incurred. One participant was so annoyed that they stated "I never, ever want to buy a tablet that doesn't allow me to place my hand on the screen." Although a decrease in accuracy was not found, participants felt very strongly that they were less accurate when holding their hand in the air ($M = 5.93$, $SD = 1.28$; $t(15) = 5.85$, $p < .001$). Many participants, however, were found to grasp the stylus much higher up on the barrel when their hand was in the air. In post-experiment follow-up, many participants were unaware of this behavior but one participant made "a conscious decision as to where (he) held along the pen, (where (his) fingers actually touched the pen, near the bottom, or the middle of the pen)." Such compensatory behaviors could have implications while interacting for extended periods.

The results showed that visual feedback, more so than resting the hand, is beneficial while targeting and writing. The hover cursor enabled participants to be 161% more accurate and 5% faster while targeting. The accuracy is likely due to participants capitalizing on the feedback for their targeting movements, and the slight increase in speed may be due to participants adapting to the latency inherent in the cursor, using their own prediction method to determine where it would end up. When feedback was unavailable, users made slower, more thoughtful movements, knowing that the lack of feedback left them to simply hope for the best. While writing, the visual feedback created a 49% improvement in accuracy and a 6% increase in duration. When the cursor was present, participants were more likely to move cautiously before tapping the screen, capitalizing on the cursor feedback for each stroke, which likely had much latency. When the cursor was not present, users were faster because they realized that no amount of aligning the stylus would increase their accuracy.

Participants reported that they noticed the hover cursor ($M = 5.88$, $SD = 1.61$; $t(15) = 4.4$, $p < .001$) and felt it made them more accurate ($M = 4.94$, $SD = 1.70$; $t(15) = 2.2$, $p < .05$), but were mixed on its usefulness for everyday tasks ($M = 4.3$, $SD = 2.12$; $t(15) = .6$, $p = 0.28$). While the hover cursor has largely been viewed as only required due to inaccuracy, 43% of participants said it would still be helpful if the stylus was perfectly accurate. With a traditional pen, the user is well aware where ink will be deposited, as the ink or lead leaves marking at those points that touch paper. On a digital device, the cursor acts as a backup feedback system, helping overcome some of the other elements that contribute to inaccuracy identified in Section 5.1.

Participants also noted that the hover cursor was helpful while initially targeting or selecting an object but was of less useful while writing because "writing is allowed to be a messy task to begin with" and "I can make corrections on the fly while writing so inaccuracy is ok". Participants also held the pen vertically during the selection task or "moved the tablet closer and looked around their hand and pen" to overcome occlusion.

5.3 Experiment 2: Influence of Nib Size

The second experiment assessed another uncommon factor, the size of the nib. As nibs get larger, they occlude more of the screen. Although one can assume the larger the nib, the less accurate one is, it is unclear how much the user experience degrades when larger nibs are used. Users can adapt to and write with many implements, but how does this extend to the digital world? Is the difference enough to warrant increased engineering efforts to reduce nibs sizes or should resources be directed elsewhere? Using identical sensing mechanisms and only modulating the nib diameter allowed an evaluation of the role of nib diameter and perception of inaccuracy.

5.3.1 Participants and Experimental Apparatus

All sixteen participants from Experiment 1 participated in Experiment 2. Half of the participants began with Experiment 1 and the other half began with Experiment 2. The experimental setup was identical to the first experiment.

5.3.2 Experimental Conditions

To assess the impact of the nib on accuracy, only one factor was manipulated, nib diameter. Although many other factors, such as nib malleability, barrel contour, and nib transparency are likely implicated, nib diameter was chosen because it is one of the most prevalent differences between active and passive styli today. It also required the least changes to the styli and ensured that the evaluation did not introduce additional inaccuracy.

As nibs strong and thin enough to fit inside the nib chamber could not be manufactured, 'caps' were 3D printed using Polylactic Acid (PLA) and glued to the end of the stylus (Figure 5.5). The printed nibs were designed such that they had the same shape and varied only in diameter. The diameters were chosen such that the largest nib replicated the nib size found with passive styli today (i.e., 5.6 mm); the remaining nib diameters were distributed within this range (i.e., 3.5 mm, and 4.4 mm). With the caps, the same activation force was necessary to trigger the stylus.

Figure 5.5. The four nibs used in the second experiment. From left to right: the 5.6 millimeter nib, the 4.4 millimeter nib, 3.5 the millimeter nib, and the 1.6 millimeter nib.

5.3.3 Tasks, Procedure, Measures, and Data Analysis

Similar to Experiment 1, participants completed the selection and writing tasks using the same target widths, target distances, and so on. Participants were also free to rest their hand on each trial and the hover cursor was visible whenever the stylus was in the hover state. With the 1.6 millimeter and 3.5 millimeter nibs, the hover cursor and targets were not occluded; with the 4.4 millimeter and 5.6 millimeter nibs, occlusion was possible depending on how the stylus was held.

The order of the four nib diameters was counterbalanced, as was the presentation of the task manipulations (i.e., target distances, rectangle width, etc.), and the tasks themselves. Similar to Experiment 1, the number of errors, the movement offset, duration, and pressure exerted were recorded and calculated for the second experiment. Participants also completed 7-point Likert scale questions and ranked the styli according to which were the most preferred and accurate.

5.3.4 Results

For the second experiment, a within-subjects repeated measures ANOVA, with Nib Diameter (levels: 1.6, 3.5, 4.4, and 5.6) as the main factor was performed for the selection and writing tasks. For the questionnaire data, one-sample t-tests were performed, comparing all responses to the neutral response (i.e., '4'), and, where appropriate, comparisons were made against the smallest or the largest nib size (i.e., '1.6 mm' and '5.6 mm'). The questionnaire responses are reported in the Discussion.

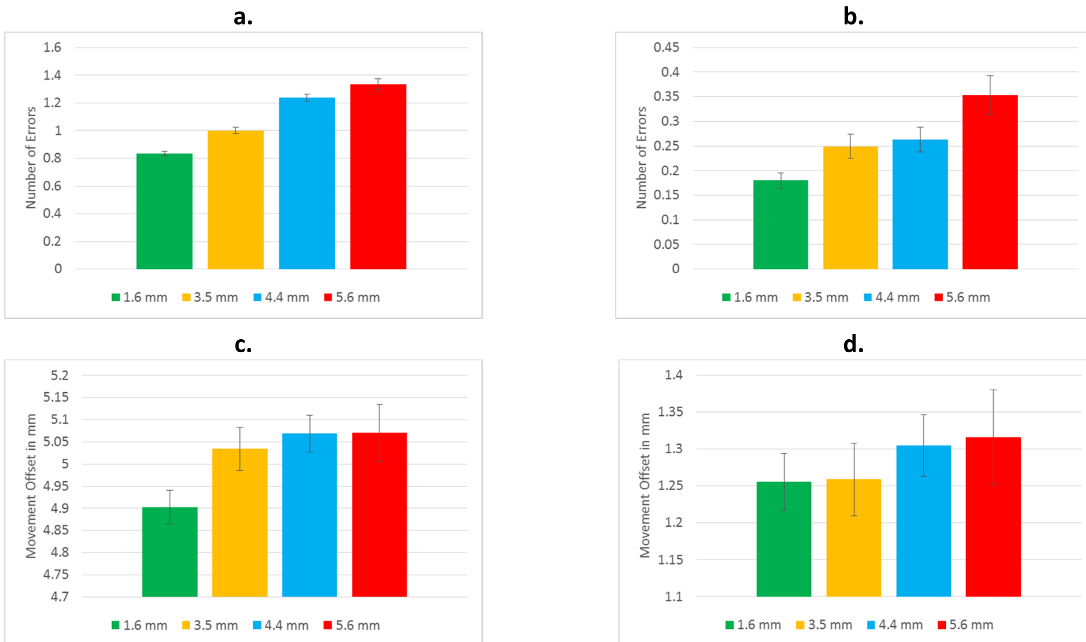


Figure 5.6. (a) The error found for the selection task, (b) the error for the writing task, (c) the movement offset for the selection task, and (d) the movement offset for the writing task. The error bars represent the standard error of the mean.

5.3.4.1 Errors

While selecting, Nib Diameter influenced the number of errors made (Figure 5.6a; $F_{3,45} = 2.9$, $p < .05$, $\eta^2 = .16$). Bonferroni-corrected paired t-tests determined that participants made significantly fewer targeting attempts with the 1.6 millimeter versus 4.4 millimeter (p < .05) and 1.6 millimeter versus 5.6 millimeter nibs (p < .01). Nib diameter thus influences accuracy, especially when diameters are atypically large (e.g., 4.4 millimeter or 5.6 millimeter). Nib Diameter did not influence the number of digits that fell outside the target rectangles (Figure 5.6b, $F_{1.3,20.75} = 0.0$, $p = .39$, $\eta^2 = .06$).

5.3.4.2 Movement Offset

The larger the nib, the less accurate were the targeting movements (Figure 5.6c; $F_{3,45} = 3.9$, $p < .05$, $\eta^2 = .21$). Bonferroni-corrected paired t-tests revealed that participants were significantly more accurate with the 1.6 millimeter compared to 3.5 millimeter nib ($p < .05$), the 1.6 millimeter versus 4.4 millimeter nib ($p < .01$), and the 1.6 millimeter versus 5.6 millimeter nib ($p < .05$). Nib Diameter is thus important when targeting, with a possible 'inaccuracy threshold' emerging at 3.5 millimeter. Nib Diameter was not found to influence the centeredness of the written digits (Figure 5.6d, $F_{1.9,29.45} = 1.6$, $p = .24$, $\eta^2 = .09$).

5.3.4.3 Pressure Exerted

In the selection task, Nib Diameter influenced the pressure exerted (Table 5.2; $F_{3,45} = 8.9$, $p < .001$, $\eta^2 = .37$). Bonferroni-corrected paired t-tests demonstrated that participants exerted significantly more force with the 3.5 millimeter versus 1.6 millimeter nib ($p < .01$), the 4.4 millimeter versus 1.6 millimeter nib ($p < .01$), and the 5.6 millimeter versus 1.6 millimeter nib ($p < .05$). Although the activation force for each nib was identical, participants may have perceived the larger nibs as less accurate, exerting more force. It thus appears that pressure is unconsciously used as a compensation mechanism for poor accuracy.

Nib Diameter also influenced the pressure exerted while writing (Table 5.2; $F_{3,45} = 12.6$, $p < .001$, $\eta^2 = .46$). Participants exerted more force with the 5.6 millimeter than 4.4 millimeter nib ($p < .001$), the 5.6 millimeter versus 3.5 millimeter nib ($p < .001$), the 5.6 millimeter nib than the 1.6 millimeter nib ($p < .01$), and the 3.5 millimeter versus 1.6 millimeter nib ($p < .01$). Similar to the selection task, it appears that the more inaccurate the stylus was perceived to be, the more force exerted.

Table 5.2. The results for the Duration and Pressure Exerted measures for Experiment 2.

		Duration in Milliseconds				Pressure Exerted			
		Selection		Writing		Selection		Writing	
		M	SEM	M	SEM	M	SEM	M	SEM
Nib Diameter	1.6 mm	883	35	7408	301	22.47	2.21	109.49	6.07
	3.5 mm	937	42	7252	340	25.98	2.06	116.98	6.18
	4.4 mm	954	36	7581	409	31.88	2.66	116.18	7.64
	5.6 mm	990	39	7788	532	33.31	3.16	134.03	7.14

5.3.4.4 Duration

In the selection task, Nib Diameter influenced the speed of writing (Table 5.2; $F_{3,45} = 5.4$, $p < .01$, $\eta^2 = .27$). Bonferroni-corrected paired t-tests demonstrated that participants were significantly faster with the 1.6 millimeter versus 4.4 millimeter nib ($p < .01$) and 1.6 millimeter versus 5.6 millimeter nib ($p < .01$). Such results align with the accuracy measures and suggest speed and accuracy disadvantages when using larger nibs. No influence of Nib Diameter was found for writing (Table 5.2; $F_{1.9,\,28.49} = 0.9$, $p = 0.41$, $\eta^2 = .06$). When combined with the accuracy results from the first experiment, it appears that nib diameter and feedback play a larger role in the speed of movement than hand posture. While inking, it also appears that feedback influences one's speed more so than nib diameter.

5.3.5 Discussion

The second experiment demonstrates how large the disparity between active and passive styli is. When using a nib similar to a passive stylus (i.e., 5.6 mm) compared to the nibs used in active styli (i.e., 1.6 mm), participants exerted 48% more pressure, exhibited a 12% increase in duration, and accuracy decreased by 60%. The use of passive styli for selection is detrimental not only in terms of accuracy, but also fatigue and the speed at which selections can be made.

When traditional accuracy measures are considered, nib diameter had little influence while writing. It is possible that the hover cursor and the feedback it provided neutralized the impact of the differing nib dimensions. Because the 4.4 millimeter and 5.6 millimeter nibs occluded the cursor and targets, this is, however, unlikely. While writing, if an initial targeting movement was incorrect, there were many opportunities for on the fly corrections. With selection, this was not possible.

Although no accuracy or duration differences were found for writing, a significant number of participants believed that the size of the stylus nib influenced, and would continue to influence, their accuracy and content neatness while inking ($M = 5.88$, $SD = 0.89$; $t(15) = 8.5$, $p < .001$). Participants felt more accurate with the smaller nibs ($M = 4.9$, $SD = 1.4$; $t(15) = 2.7$, $p < .01$). A significant number of participants (i.e., 63 percent) ranked the 1.6 millimeter nib as the most preferred ($t(15) = 2.8$, $p < .01$) and accurate ($t(15) = 2.7$, $p < .01$) whereas 75% felt the 5.6 millimeter nib was the least preferred ($t(15) = -2.0$, $p < .05$) and 63%, the least accurate ($t(15) = -2.6$, $p < .05$). Even though they were no more accurate, the perception of accuracy plays an important role in the conscious decisions and unconscious movements made. When asked if a nib thinner than 1.6 millimeters was desired, none of the participants were receptive to this idea. Although many indicated that they used very precise pens and pencils on a daily basis (i.e., 0.7 mm), they had different expectations for digital devices. It thus remains to be seen how closely the design of digital experiences should mimic physical experiences.

Those who preferred the 3.5 millimeter and 4.4 millimeter nibs indicated that the 1.6 millimeter nib was too small (e.g., "the 3.5 and 4.4 millimeter pens were similar to the markers I use and like") and professed that the 3.5 millimeter and 4.4 millimeter nibs looked identical in size to them. This visual similarity may explain the plateauing found with the pressure exerted. As only the perceived accuracy was manipulated, participants unable to discern between the 3.5 millimeter and 4.4 millimeter nibs assumed that they would produce the same level of inaccuracy, thus they applied the same compensatory pressure.

5.4 Overall Discussion

Both experiments demonstrated that nib size, visual feedback, and the ability to rest the hand play roles in the speed, accuracy, and pressure with which one selects targets and writes. The same participants and tasks were used in the two experiments, hence a comparison between the effect sizes for the three conditions can be made (Table 5.3). The feedback provided to the user from the hover cursor thus appears to be of highest importance to accuracy. When the cursor was not present or difficult to see (due to nib occlusion), participants were much less accurate and moved slower.

Table 5.3. The resulting effect sizes for the Hover Cursor, Hand Resting, and Nib Diameter conditions. (-) denotes the measure was not found to be significant.

		Hover Cursor η^2	Hand Resting η^2	Nib Diameter η^2
Selection	Error	0.38	-	0.16
	Movement Offset	0.49	0.03	0.21
	Pressure Exerted	-	0.23	0.37
	Duration	0.07	-	0.27
Writing	Error	0.18	-	-
	Movement Offset	0.50	-	-
	Pressure Exerted	-	0.29	0.46
	Duration	0.10	0.27	-

Although the hover cursor is a purely digital construct, it is incredibly useful. Given the role that the hover cursor plays in accuracy, efforts should be made to increase the detectable height of the stylus above the screen and advance hardware such that the hover state exists with capacitive devices. The sooner users receive feedback about the location of the stylus, the better their movements. Given that users complain about inaccuracy and legibility while writing and selecting, it would be fruitful to explore extending precision selection widgets such as the Bubble Cursor (Grossman & Balakrishnan, 2005) or Pointing Lenses (Ramos et al., 2007) for use in sketching and writing scenarios.

Nib diameter was found to be less important than visual feedback, but still more important than hand posture. Although larger nibs lead to less accuracy while selecting targets and more pressure while selecting and writing, nib diameter does not influence accuracy in a linear manner. There could be 'accuracy plateaus', wherein decreases in accuracy and increases in pressure only accompany visually indistinguishable nib diameters. As only four nib sizes were tested, this cannot be confirmed definitively. The diameter results also introduce interesting implications and motivation for multi-pen systems or styli with interchangeable, multi-surface, or chiseled nibs. Will users unintentionally apply different pressure while using different sides of the stylus? Should this behavior be compensated for or harnessed for contextual information or intent? The exploration into stylus aesthetics was confined to nib diameter, but many other facets such as nib deformity, nib shape, and barrel tapering should also be explored.

Although the results favored visual feedback and nib design, concerns over fatigue and comfort should not be discounted. When asked if participants would prefer i) an accurate pen but were prevented from resting their hand or ii) an inaccurate pen but were allowed to rest their hand, 63% of participants favored hand resting. Surprisingly, users are willing to accept inaccuracy if it means they can interact naturally. In the experiment, each task lasted a few minutes and participants were free to rest their elbow on the table. While using a tablet in one's lap, holding it in the air, or interacting for longer durations, fatigue will eventually set in and led to less controlled movements and to some degree of inaccuracy. Understanding the fatigue/inaccuracy relationship is thus important and provides continued motivation to solve unintended touch and develop surface textures that provide appropriate sensory feedback and support smooth, dragging hand motions.

Pressure has commonly been used in the design of pressure-sensitive user interface elements or to beautify ink (Ramos, Boulos, & Balakrishnan, 2004; Zitnick, 2013). Determining the pressure that is applied when nib size and hand resting changes, leads to interesting questions for future research. Should developers strive for consistent ink rendering when different styli are used? Should compensatory adjustments be made when the hand is in the air or when the hand is resting and increased pressure is likely to be applied? As pressure appears to used as a compensatory adaptation, it is important to consider the implications of pressure beyond accuracy (e.g., use of pressure for unintended touch, biometrics, and so on).

The present work demonstrates how diverse and complex the user's experience of accuracy is within the tablet ecosystem today. There is thus still much room for innovation and exploration, especially when one considers performance, speed, and comfort. Inaccuracy can no longer be thought of as a simple by-product of poor hardware decisions or hardware limitations. Practitioners and manufacturers must consider how the technology will support natural user behaviors, and researchers and developers should capitalize on user behavior and design software to compensate for hardware limitations. Now, more so than ever, cross-device and cross-stylus designers and developers need to consider the how to bridge the breadth and depth dichotomy within the tablet ecosystem. Does one design for the most basic stylus features or have disparate experiences across devices and styli? Should the same compensatory UI elements be provided across all devices? The new Pencil (Fifty Three, Inc. 2014) and Ink (Adobe, 2014) auxiliary styli offering pressure sensing and the beginnings of solutions to unintended touch are a good start at bridging this gap, but much work is still needed to provide a coherent user experience across the ecosystem.

5.5 Summary

Inaccuracy has long been a complaint and frustration of users while interacting with a stylus. It is often assumed that inaccuracy is only caused by visual parallax or nib occlusion. In the present work, it was determined that many factors influence accuracy on tablets today. While advancements to the underlying stylus technology should continue, the present exploration has identified that much more attention needs to be devoted to the other two categories, the design of the stylus and user interaction.

Using ISO 9241 selection and inking tasks, an evaluation of a subset of these factors, namely stylus design, the feedback provided by the hover cursor, and the users' ability to test their hand was performed. The experiments found that the feedback provided by the hover cursor is of upmost importance to accuracy. Nib diameter was also demonstrated to affect accuracy, with users' perceptions of a stylus influencing the pressure they exerted and the accuracy measured. The ability to rest one's hand on the screen was found to play little role in accuracy. The present experiments have provided rich information regarding the factors that influence inaccuracy and lead to the disparities found in the stylus and tablet ecosystem today.

Chapter 6

Conclusions

The work presented in this thesis focused on observing the stylus and tablet experience today in an attempt to i) identify the problems with today's devices, ii) develop a deep understanding of the behavioral and performance issues implicated in each issue, and iii) explore solutions to prevent them. Chapter 2 presented an exploratory study that uncovered many of the issues and behavioral adaptions found with styli and tablets today. By comparing the digital world to the 'gold standard' of pen and paper, the study identified five factors (i.e., device latency, unintended touch, stylus accuracy, stylus and device aesthetics, and stroke beautification) believed to be the fundamental causes of decreased performance and satisfaction today. Three of these factors were identified as being significantly more impactful, and thus directed and guided the rest of the research presented in this thesis.

As little to no focus has been placed on understanding the role of latency while inking with a stylus, psychophysical just-noticeable difference experiments were conducted using a prototype low latency, High Performance Stylus System. It was found that participants could discriminate between very low levels of latency while dragging and scribbling (i.e., ~1 versus 2 milliseconds, and ~7 versus 40 milliseconds, respectively), latency perception while inking was worse (~50 milliseconds) than while performing non-inking tasks (~2-7 milliseconds), and that latency perception is not based on the distance from the stylus' nib to the ink, but rather on the presence of a visual referent, such as the hand or stylus. The present exploration, in additional to a synthesis of prior work on touch-based latency, informed the creation of the *Latency Perception Model*, a framework upon which latency knowledge can grow and the underlying mechanisms of perception can be further understood and explored.

During the exploratory experiment, devices unable to distinguish between *intended touch*, i.e., interaction on the screen intended for action, and *unintended touch*, i.e., incidental interaction from the palm, forearm, or fingers, resulted in stray ink strokes and accidental navigation, frustrating users. A data collection experiment was thus conducted wherein participants performed inking tasks. Natural tablet and stylus behaviors were observed and analyzed from both digitizer and behavioral perspectives. An analysis and comparison of novel and existing unintended touch algorithms revealed that the use of larger rejection regions and the use of stylus information to dictate these regions can greatly reduce problems with unintended touch. The analysis also identified many natural stylus behaviors that undermine some of the potential solutions to unintended touch that were examined and underscored the importance of being mindful of application and ecosystem demands.

While much work has focused on the effects of visual parallax and occlusion, there is still much unknown about the role of the user and the stylus when it comes to accuracy. The present work enhanced this understanding by identifying ten factors that can influence the accuracy of stylus input on tablets. Further, the results of two user studies that systematically evaluated three of the lesser understood, but increasingly important, factors (i.e., hand posture, visual feedback, and nib design) were reported. The results determined that the presence of visual feedback and the dimensions of the stylus nib were crucial to the accuracy attained and pressure exerted with the stylus. The ability to rest one's hand on the screen, while providing comfort and stability, had surprisingly little influence on the accuracy experience.

Before devices of the future can integrate novel features, it is imperative that developers and designers create experiences that meet users' expectations. Across all of the experiments presented in the thesis, an enriched understanding of the stylus experience, including its grievances and pitfalls, was brought to light. The identification of the specific elements that influence the stylus experience, i.e., device latency, stylus accuracy, and unintended touch, along with a detailed experimental analysis of each factor should help guide the development of new styli and tablets and ensure that satisfying user experiences are had for years to come.

6.1 Thesis Contributions

The work described in this thesis has made several contributions to the field of Human-Computer Interaction, specifically in the fields of pen and touch interaction and pen computing:

Chapter 2

- **Set of reproducible tasks and measures:** The use of widely available devices and reproducible measurement techniques allowed for a comparison of the physical and digital worlds and should enable other researchers to perform their own evaluations and assess the suitability of other devices and input implements for inking.
- **Comparison of behavioural and performance differences between digital and analog media:** Although many have assumed that the digital experiences should mimic the physical experience, little is known

about the physical experience. Performing an explicated comparison between the two has provided immense information regarding natural interaction, common writing sizes, and user preferences and expectations.

- **Identification of behavioural adaptations, hand movements, and grips unique to digital devices:** This chapter provided the first quantitative evidence of the adaptions and unique behaviors displayed by users (including left handers) when using digital devices. The identification of such behaviors will not only help in the development of solutions to overcome such issues such as unintended touch, but also provide insight into how such behavior could be harnessed in the future for novel interaction and functionality.
- **Identification and prioritization of outstanding issues prohibiting satisfying stylus experiences:** Identifying the importance of latency, accuracy, unintended touch, stylus and device aesthetics, and stroke beautification to users and their behavior has provided a foundational focus to which much future work and attention can be devoted. It additionally inspired the studies and results presented in Chapters 3, 4, and 5.

Chapter 3

- **Identification of the minimum perceivable latency while performing stylus-based tasks**: Using a prototype High Performance Stylus System, users' perception of latency was investigated using a variety of stylus-based tasks. The ability to test a range of delays, many of which are not achievable in the near future, provided a wealth of information and understanding regarding latency perception.
- **Understanding of the factors implicated in latency perception**: The results detailed in this chapter uncovered the importance of task demands and the strategies employed to discriminate latency. It was found that while make latency judgements, the relative motion of the hand or stylus to the ink was used instead of the distance between the ink and stylus and that the presence of the stylus and hand are necessary for latency perception.
- **Latency Perception Model**: The results from the three experiments detailed in this chapter provided deep insight into how latency is perceived and formed a generalized model and blueprint for future explorations into latency perception.

Chapter 4

- **Technical limitations and requirements governing unintended touch:** This chapter presents the first specification and detailed understanding of the limitations inherent in performing unintended touch along different segments of the data pipeline. From the identification of the timing limitations to the types of touch information that need to be rejected, such specifications will greatly help the research community to focus and continue addressing the problem of unintended touch.

- **Evaluation and comparisons of approaches to unintended touch given present and future technological limitations:** The thorough approach taken to understand potential unintended touch solutions revealed not only the benefits of using different sources of data but also the delicate balance that exists between perfect rejection and natural interaction. It also provides some of the first data on common touch input locations (including from left handed users), natural nib heights while gesturing, and the time necessary for finger-based touch input to be distinguished from other touch-based input.
- **Fruitful avenues for future research:** The results of this exploration and the behavioral patterns identified have been synthesized into two broad areas for future work and technical innovation (i.e., the use of additional data streams and software and hardware innovations) that should be of great interest to the pen computing community and stimulate the stylus ecosystem for many years to come.

Chapter 5

- **Identification of the factors influencing accuracy:** The work presented in this chapter provides the first in-depth analysis of the factors influencing accuracy. As occlusion and visual parallax have long been regarded as the only factors, the identification and classification of eight other factors is of great importance of pen computing.
- **Systematic evaluation of a subset of these factors:** By systematically evaluating factors using standardized tasks and measures, a richer understanding of accuracy was attained, extending past traditional factors that have been the focus as of late.
- **Relative importance of factors influencing accuracy:** The quantitative results and participant comments from within this chapter underscored the importance of visual feedback and nib aesthetics to accuracy, while surprisingly devaluing the importance of supporting natural hand postures.

6.2 Future Work

Although the present work has made many contributions to pen computing, there are still many questions that remain unanswered and ripe for exploration. Chapter 2 identified the importance of device aesthetics and stroke beautification, but little attention was given to these factors that will become increasingly important once accuracy, latency, and unintended touch no longer plague pen computing. Understanding how users expect strokes to look, especially given the visual appearance of a stylus, will be crucial as the variety of implements compatible with tablets increases, i.e., should a 'marker-style' pen be able to draw brush and pencil strokes? With traditional pencils, repeated use leads to wear and subsequently different markings left by the lead. Determining if this natural decaying process is useful and how it could be integrated within the digital world, will greatly increase the relationship users have to their content. Armed with advances in machine learning and graphics, the developers and designers of the tablets of tomorrow will need to be mindful of such possibilities.

There still remain many questions about the relationship between accuracy and texture, especially when one considers the variety of nib and surface textures possible. With advances in material science and novel fabrication techniques, there will be a plethora of materials available to construct new stylus nibs and surface textures. Determining the relation between various textures, the sensations felt by the user and their resulting accuracy should provide many benefits and much knowledge for human-computer interaction and psychology. As users continue to desire and expect digital experiences that mimic the physical world, it will be important to understand how user experience and performance change when displays are able to dynamically change texture. A more immersive, natural experience that is likely if displays could simulate canvas while painting or loose-leaf while writing.

Much foundational knowledge and many behavioral observations have been presented throughout Chapters 2-5. Many of the behavioral observations should be of great value to the research community and provide context for many of the novel interaction techniques and stylus-based functionality in the future. Although this thesis has laid a foundation in terms of unintended touch, accuracy, and latency, there is still much room for additional exploration. One area that did not receive much attention was the relationship between the various factors. The systematic evaluation that was undertaken has allowed for a deeper understanding of each factor, but future work should extended upon this to explore how unintended touch is improved when latency is decreased, how decreased latency influences the use of the hover state (to improve accuracy), and the relative importance and acceptance of the factors when taken together. The thesis also put forth introductory hypotheses regarding the underlying biological and psychology processes that govern interaction. Further empirical examination of these hypotheses is required, as is an understanding of how the visual feedback and the motor systems influence one's perception with a stylus-enabled system. Employing eye-tracking technology can also allow for a unique perspective into the user-based factors involved in (perceived) accuracy and potentially uncover new techniques or methods to alleviate inaccuracy.

As alluded to in Chapter 5, there is still much work to be done within the stylus ecosystem to create cohesive user experiences. Due to competing corporate interests, there will always be diversity in the marketplace, with less expensive, limited styli-enabled devices and higher-end, fully-functional devices. Within the context of pen computing, the 'breadth versus depth' dichotomy has become ever more prevalent. Such diversity will require designers to think carefully about the transition between active and passive styli, the validity of differentiated passive styli, and stylus-enabled phone and tablet experiences. As auxiliary styli become more prevalent, determining how to best support their varying features across devices or applications will be imperative to providing a cohesive experience. Much research and innovation is required to develop appropriate feedback methods and visualization techniques to increase awareness about currently permitted, restricted, and alternative functionalities. Additionally, increased attention should be given to adaptive interface widgets and interfaces that automatically transition between input states and stylus functionality. From a more holistic viewpoint, the metaphors assigned to the finger and stylus should be better defined, consistent across platforms and devices, and the appropriateness of each input modality should better understood (taking into consideration device limitations and multi-user scenarios). The introduction of modular devices composed of interchangeable input and output modules, such as phones, watches, and likely tablets, will only complicate these issues further.

References

Adobe. (n.d.). Adobe Ink & Slide. Retrieved June 2014 from http://www.adobe.com/products/ink-and-slide.html.

Adelstein, B.D., Begault, D.R., Anderson, M.R., & Wenzel, E.M. (2003). Sensitivity to Haptic-Audio Asynchrony. In the *Proceedings of International Conference on Multimodal Interfaces*, 73-76.

Admin. (2013). Sharp Released the Pencil Handwriting Touchscreen, May Be a Weapon of the iPhone 6. Retrieved From http://www.ztechnews.com/sharp-released-the-pencil-handwriting-touch-screen-may-be-a-weapon-of-the-iphone-6/.

Aitenbichler, E. & Schnelle-Walka, D. (2010). An Extensible Architecture for Multi-Touch and Pen Interactive Tablets. In the *Proceedings of Engineering Interactive Computing Systems*.

Alamargot, D., Chesnet, D., Dansac, C., & Ros, C. (2006). Eye and Pen: a New Device for Studying Reading during Writing. *Behaviour Research Methods*, *38(2),* 287-299.

Anderson, G., Doherty, R., & Ganapathy, S. (2011). User Perception of Touch Screen Latency. *Design, User Experience, & Usability Part 1*, 195-202.

Annett, M., Anderson, F., Bischof, W.F., & Gupta, A. (2014). The Pen Is Mightier: Understanding Stylus Behavior While Inking on Tablets. In the *Proceedings of Graphics Interface,* 193-200.

Annett, M., Ng, A., Dietz, P., Gupta, A., Bischof, W.F. (2014). How Low Should We Go? Understanding the Perception of Latency While Inking. In the *Proceedings of Graphics Interface*, 167-174.

Alcantara, T., Ferreira, J., & Maurer, F. (2013). Interactive Prototyping of Tabletop and Surface Applications. In the *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 229-238.

Atmel. (n.d.). Revolutionizing the Touch User Experience. Retrieved June 2014 from http://www.atmel.com/microsite/stylus/default.aspx.

Badam, S.K., Chandrasegaran, S., Elmqvist, N., & Ramani, K. (2014). Tracing and Sketching Performance Using Blunt-Tipped Styli on Direct-Touch Tablets. In the *Proceedings of Advanced Visual Interfaces,* 193-200.

Balakrishnan, R. & Hinckley, K. (1999). The Role of Kinesthetic Reference Frames in Two-Handed Input Performance**.** In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 171-178.

Bargeron, D. & Moscovich, T. (2003). Reflowing Digital Ink Annotations. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 385-392.

Bernstein, N. (1967). *The Coordination and Regulation of Movements.* Oxford: Pergamon Press.

Bi, X., Moscovich, T., Ramos, G., Balakrishnan, R., & Hinckley, K. (2008). An Exploration of Pen Rolling for Pen-Based Interaction. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 191-200.

Blanch, R., Guiard, Y., & Beaudouin-Lafon, M. (2004). Semantic Pointing: Improving Target Acquisition with Control-Display Ratio Adaptation. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 519-526.

Brandl, P., Forlines, C., Wigdor, D., Haller, M., & Shen, C. (2008). Combining and Measuring the Benefits Tabletops and Surfaces of Bimanual Pen and Direct-Touch Interaction on Horizontal Interfaces. In the *Proceedings of the Working Conference on Advanced Visual Interfaces*, 154-161.

Brandl, P., Leitner, J., Seifried, T., Haller, M., Doray, B., & To, P. (2009). Occlusion-Aware Menu Design for Digital Tabletops. In the *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*, 3223-3228.

Breuel, T.M., Janssen, W.C., Popat, K., & Baird, H.S. (2002). Reflowable Document Images. *Web Document Analysis: Challenges and Opportunitie*s, Chapter 1.

Briggs, R.O., Dennis, A.R., Beck, B.S., & Nunamaker, J.F. (1992). Whither the Pen-Based Interface? *Journal of Management Information Systems*, 71-90.

Brush, A.J., Barageron, D., Gupta, A., & Cadiz, J.J. (2001). Robust Annotation Positioning in Digital Documents. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 285-292.

Butter, A. & Pogue, D. (2009). *Piloting Palm*. Retrieved From http://www.oocities.org/firstspeaker.geo/books/nonfiction/piloting_palm.html.

Buxton, W.S. (1990). A Three-State Model of Graphical Input. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 449-456.

Carlton, L.G. (1981). Visual Information: the Control of Aiming Movements. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *33(1),* 87-93.

Cao, X., Wilson, A.D., Balakrishnan, R., Hinckley, K., & Hudson, S.E. (2008). ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. In the *Proceedings of the 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*, 129-136.

Cockburn, A., Ahlstrom, D., & Gutwin, C. (2008). Understanding Performance in Touch Selections: Tap, Drag, & Radial Pointing Drag With Finger, Stylus, & Mouse. *Journal of Human-Computer Studies*, *70*, 218-233.

Coldewey, D. (2011). Synaptics New Touchscreens Can Detect the Head of a Pin. Retrieved from http://techcrunch.com/2011/02/15/synaptics-new-touchscreens-can-detect-the-head-of-a-pin/.

Coen-Calgi, R.C., Coraggio, P., Napoletano, P., Schwartz, O., Ferraro, M., & Boccignone, G. (2009). Visuomotor Characterization of Eye Movements in a Drawing Task. *Vision Research*, *49(8),* 810-818.

Dahl, S. & Bresin, R. (2001) Is the Player More Influenced by the Auditory than the Tactile Feedback from the Instrument? In the *Proceedings of the COST-Workshop on Digital Audio Effects,* 194-197.

Dang, C.T., Straub, M., & André, E. (2009). Hand Distinction for Multi-Touch Tabletop Interaction. In the *Proceedings of the ACM International Conference on ACM International Conference on Interactive Tabletops and Surfaces*, 101-108.

Davis, M.R. & Ellis, T.O. (1964). The RAND Tablet: a Man-Machine Graphical Communication Device. *Memorandum RM-4122-ARPA*, 1.

Dreyfuss, H. & Tilley, A. (2002). The *Measure of Man & Woman*. John Wiley & Sons, Inc., New York.

Ellis, T.O., Heafner, J.F., & Sibley, W.L. (1969). The GRAIL Language and Operations. *Memorandum RM-6001-ARPA*, 3.

EO Personal Communicator (n.d.) In *Wikipedia*, Retrieved June 2014 from http://en.wikipedia.org/wiki/eo_personal_communicator.

Ewerling, P., Kulik, A., & Froehlich, B. (2012). Finger and Hand Detection for Multi-Touch Interfaces Based on Maximally Stable Extremal Regions. In the *Proceedings of the ACM International Conference on ACM International Conference on Interactive Tabletops and Surfaces*, 173-182.

Fifty Three, Inc. (n.d.). In Wikipedia. Retrieved June 2014 from http://en.wikipedia.org/wiki/fiftythree,_inc.

Font Shop. (2010). *Typeface Anatomy*. Retrieved June 2014 from http://www.fontshop.com/education/pdf/typeface_anatomy.pdf.

Fayol, M., Alamargot, D., & Berninger, V.W. (2012). *Translation of Thought to Written Text While Composing: Advancing Theory, Knowledge, Research Methods, Tools, & Application*. Psychology Press, Taylor & Francis Group, New York.

Fekete, J.D., Bizouarn, É., Cournarie, É., Galas, T., & Taillefer, F. (1995). Tictactoon: a Paperless System for Professional 2D Animation. In the *Proceedings of SIGGRAPH*, 79-95.

Fitzmaurice, G. Balakrishnan,R., Kurtenbach, G., & Buxton, B. (1999). An Exploration into Supporting Artwork Orientation in the User Interface. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 167-174.

Forlines, C. & Balakrishnan, R. (2008). Evaluating Tactile Feedback and Direct Vs. Indirect Stylus Input in Pointing and Crossing Tasks. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1563-1572.

Freed, A. Chaudhary, A. & Davila, B. (1997). Operating System Latency Measurement and Analysis for Sound Synthesis and Processing Applications. In the *Proceedings of International Computer Music Conference*, 479-481.

Giachis, C.D., l'Anson, S., & Prytherch, D. (2001). Haptic Discrimination of Different Types of Pencils during Writing. *Ergonomics*, *54(8),* 684-689.

Gerken, J., Jetter, H.C., Schmidt, T., & Reiterer, H. (2010). Can "Touch" Get Annoying? In the *Proceedings of the ACM International Conference on ACM International Conference on Interactive Tabletops and Surfaces*, 257-258.

Goonetilleke, R.S., Hoffman, E.R., & Luximon, A. (2009). Effects of Pen Design on Drawing and Writing Performance. *Applied Ergonomics*, *40*, 292-301.

Gowen, E. & Miall, R.C. (2006). Eye-Tracking Interactions in Tracing and Drawing Tasks. *Human Movement Science*, 25, 568-585.

Gridpad. (n.d.). In *Wikipedia*. Retrieved June 2014 from http://en.wikipedia.org/wiki/gridpad.

Grossman, T. & Balakrishnan, R. (2005). The Bubble Cursor: Enhancing Target Acquisition by Dynamic Resizing of the Cursor's Activation Area. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 281-290.

Grossman, T., Hinckley, K., Baudisch, P., Agrawala, M., and Balakrishnan, R. (2006). Hover Widgets: Using the Tracking State to Extend the Capabilities of Pen-Operated Devices. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 861-870.

Guiard, Y. (1987). Asymmetric Transfer of Labor in Human Skilled Bimanual Action: the Kinematic Chain as a Model. *Journal of Motor Behavior*, *19*, 486-517.

Guiard, Y. & Millerat, F. (1984). Writing Postures in Left-Handers: Inverters are Hand-crossers. *Neuropsychologia*, 22(4), 535-538.

Gurley, B.M. & Woodward, C.E. (1959). Light-Pen Links Computer to Operator. *Electronics*, 85-87.

Hancock, M. & Booth, K. (2004). Improving Menu Placement Strategies for Pen Input. In the *Proceedings of Graphics Interface*, 221-230.

Henzen, A., Ailenci, N., Di Fiore, F., Van Reeth, F., & Patterson, J. (2005). Sketching With a Low-Latency Electronic Ink Drawing Tablet. In the *Proceedings of 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, 51-60.

Henzen, A., Ailenci, N., Di Fiore, F., Van Reeth, F., Vansichem, G., Zehner, R.W., & Amundson, K. (2004). An Electronic Ink Low Latency Drawing Tablet. *Society for Information Display Symposium Digest of Technical Papers*, *35(1)*, 1070-1073.

Hinckley, K. & Wigdor, D. (2012). Input Technologies and Techniques. In J. Jacko (Ed.), *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Third Edition* (pp. 99-132). CRC Press.

Hinckley, K., Chen, X., and Benko, H., (2013). Motion and Context Sensing Techniques for Pen Computing. In the *Proceedings of Graphics Interfac*e, 71-78.

Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., & Buxton, B. (2010). Pen + Touch = New Tools. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 27-36.

Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., Buxton, B. (2010). Manual Deskterity: An Exploration of Simultaneous Pen + Touch Direct Input. In the *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*, 2793-2802.

Holz, C. & Baudisch, P. (2011). Understanding Touch. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 2501-2510.

Holzinger, A., Holler, M., Schedlbauer, M., & Urlesberger, B. (2008). An Investigation of Finger versus Stylus Input in Medical Scenarios. In the *Proceedings of the International Conference on Information Technology Interfaces*, 433-438.

Hout, S., Nancel, M., & Beaudouin-Lafon, M. (2008). Push Menu: Extending Marking Menus for Pressure-Enabled Input Devices. *Inria Technical Report*.

ISO/TS 9241-9:2000E. (2000). *Ergonomic Requirements for Office Work With Visual Display Terminals (Vdts)-Part 9: Requirements for Non-Keyboard Input Devices*. International Organization for Standardization.

ISO/TS 9241-411:2012(E). (2012). *Ergonomics of Human-System Interaction-Part 411: Evaluation Methods for the Design of Physical Input Devices*. International Organization for Standardization.

Jota, R., Ng, A., Dietz, P., & Wigdor, D. (2013). How Fast Is Fast Enough? A Study of the Effects of Latency in Direct-Touch Pointing Tasks. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 2291-2300.

Kay, A. (1969). *The Reactive Engine*. (Doctoral Dissertation). University of Utah, Salt Lake City, Utah, United States of America.

Kyung, K. & Lee, J. (2008). Wubi-Pen: Windows Graphical User Interface Interacting With Haptic Feedback Stylus. In the *Proceedings of SIGGRAPH New Tech Demos,* 42.

Kyung, K. & Park, J. (2007). Ubi-Pen: Development of a Compact Tactile Display Module and Its Application to a Haptic Stylus. In the *Proceedings of the Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 109-114.

Kyung, K., Lee, J.Y., Park, J. (2007). Pen-Like Haptic Interface and Its Application on Touch Screen. In the *Proceedings of the IEEE Interactional Conference on Robot & Human Interactive Communication*, 9-13.

Laundry, B. (2011). *Sheet Music Unbound: a Fluid Approach to Sheet Music Display and Annotation on a Multi-Touch Screen.* (Master's Thesis). University of Waikato, Hamilton, New Zealand.

Lawler, R. (2013). *Sharp's 7-Inch AQUOS Pad SHT21 Brings Low-Power IGZO LCD Tech to Tablets in Japan*. Retrieved from http://www.engadget.com/2012/10/17/sharp-aquos-pad-sht21-igzo-lcd/.

Lee. J.C., Dietz, P., Leigh, D., Yerazunis, W.S., Hudson, S. (2004). Haptic Pen: Tactile Feedback Stylus for Touch Screens, In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies,* 291-294.

Lee, J.C., Dietz, P., Maynes-Aminzade, D., Raskar, R., & Hudson S. (2004). Automatic Projector Calibration with Embedded Light Sensors. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 123-126.

Lee, J.C., Hudson, S.E., Summet, J.W., & Dietz, P. (2005). Moveable Interactive Projected Displays Using Projector Based Tracking. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 63-74.

Lee, D., Son, K., Lee, J.H., and Bae, S. (2012). PhantomPen: Virtualization of Pen Head for Digital Drawing Free From Pen Occlusion & Visual Parallax. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies,* 331-340

Leitner, J., Powell, J., Brandl, P., Seifried, T., Daller, M., Dorray, B., & To, P. (2009). Flux: a Tilting Multi-Touch and Pen-Based Surface. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3211-3216.

Levy, J. & Reid, M. (1978). Variations in Cerebral Organization as a Function of Handedness, Hand Posture in Writing, and Sex. *Journal of Experimental Psychology: General, 107(2)*, 119-144.

Liang, R., Chang, K., Su, C., Weng, C., Chen, B., & Yang, D. (2012). GaussSense: Attachable Stylus Sensing Using Magnetic Sensor Grid. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 319-326.

Liang, R., Su, C., Weng, C., Cheng, B., Chen, B., & Yang, D. (2012). GaussBrush: Drawing With Magnetic Stylus. In the *Proceedings of SIGGRAPH Asia Emerging Technologies*, 11.

Lim, C. (2003). An Insight into the Freedom of Using a Pen: Pen-Based System and Pen-And-Paper. In the *Proceedings of 6th Asian Design International Conference*, 384-391.

Lin, C.L. Li, C.S., Chang, Y.M., Lin, T.C., Chen, J.F., & Lin, U.C. (2013). Pressure Sensitive Stylus and Algorithm for Touchscreen Panel. *Journal of Display Technology*, *9(1),* 17-23.

Liwicki, M., Rostanin, O., El-Neklawy, S., & Dengel, A. (2010). Touch & Write: a Multi-Touch Table With Pen-Input. In the *Proceedings of the 9th International Workshop on Document Analysis Systems*, 479-484.

Lopes, P., Mendes, D., Araujo, B., & Jorge, J.A. (2001). Combining Bimanual Manipulation and Pen-Based Input for 3-D Modeling. In the *Proceedings of Eurographics Symposium on Sketch-Based Interfaces and Modeling*, 15-22.

Lu, Y., Yu, F., Finkelstein, A., & Diverdi, S. (2012). Helping Hand: Example-Based Stroke Stylization. *ACM Transactions on Graphics*, *31(4),* 1-10.

Mack, R. & Lang, K. (1989). Benchmark Comparison of Mouse and Touch Interface Techniques for an Intelligent Workstation Windowing Environment. *Human Factors and Ergonomics Society Annual Meeting*, 325-329.

Mackenzie, I.S., Kauppinen, T., & Silfverberg, M. (2001). Accuracy Measures for Evaluating Computer Pointing Devices. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 9-16.

Mackenzie, I.S., Sellen, A., & Buxton, B. (1991). A Comparison of Input Devices In Elemental Pointing and Dragging Tasks. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 161-166.

Mackenzie, I.S. & Ware, C. (1993). Lag As a Determinant of Human Performance in Interactive Systems. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 488-493.

Maki-Patola, T. (2005). Musical Effects of Latency. *Suomen Musiikintutkijoiden, 9,* 82-85.

Maki-Patola, T. & Hamalainen, P. (2004). Latency Tolerance for Gesture Controlled Continuous Sound Instrument without Tactile Feedback. In the *Proceedings of International Computer Music Conference*, 1-5.

Matulic, F. & Norrie, M.C. (2012). Empirical Evaluation of Uni- and Bimodal Pen and Touch Interaction Properties on Digital Tabletops. In the *Proceedings of ACM International Conference on Interactive Tabletops and Surfaces*, 143-152.

Mcmillian, N.A. & Creelman, D.C. (1991). *Detection Theory: A User's Guide*, Cambridge University Press, Cambridge.

Miall, R.C., Imamizu, H., & Miyauchi, S. (2000). Activation of the Cerebellum in Co-Ordinated Eye and Hand Tracking Movements: An FRMI Study. *Experimental Brain Research*, *135*, 22-33.

Microsoft. (2013). Microsoft Grows Surface Family. Retrieved from http://www.microsoft.com/en-us/news/press/2013/jan13/01-22surfacefamilypr.aspx.

Microsoft. (2014). Windows Hardware Certification Requirements. Retrieved from http://msdn.microsoft.com/en-us/library/windows/hardware/dn423132.

Miller, R.B. (1968). Response Time in Man-Computer Conversational Transactions. *Fall Joint Computer Conference*, 267-277.

Mizobuchi, S., Terasaki, S., Keski-Jaskari, T., Nousiainen, J., Ryynanen, M., & Silfverberg, M. (2005). Making an Impression: Force-Controlled Pen Input for Handheld Devices. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1661-1664.

Montag, M., Sullivan, S., Dickey, S., & Leider, C. (2011). A Low-Cost, Low-Latency Multi-Touch Table with Haptic Feedback for Musical Applications. In the *Proceedings of the International Conference on New Interfaces for Musical Expression*, 8-13.

Mohr, A., Xu, D.Y., & Read, J. (2010). Evaluation of Digital Drawing Devices with Primary School Children - A Pilot Study. In the *Proceedings of the Interactive Conference on Computer-Aided Learning*, 830-833.

Muller, H., Gove, J.L., & Webb, J.S. (2012). Understanding Tablet Use: a Multi-Method Exploration. In the *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Service*, 1-10.

Murugappan, S., Vinayak, Elmqvist, N., and Ramani, K. (2012). Extended Multitouch: Recovering Touch Posture and Differentiating Users Using a Depth Camera. In the *Proceedings of the ACM Symposium on User Interfaces and Software and Technology*, 487-496.

Newton (Platform). (n.d.). In *Wikipedia*. Retrieved June 2014 from http://en.wikipedia.org/wiki/newton_(platform).

Ng, A., Annett, M., Dietz, P., Gupta, A., Bischof, W.F. (2014). In the Blink of an Eye: Investigating Latency Perception during Stylus Interaction. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1103-1112.

Ng, A., Lepinski, J., Wigdor, D., Sanders, S., & Dietz, P. (2012). Designing for Low-Latency Direct-Touch Input. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies,* 453-462.

Openframeworks. (n.d.). *Ofxtouch Is a Multitouch Add-On for Openframeworks*. Retrieved June 2013 from https://code.google.com/p/ofxtouch/.

Oldfield, R.C. (1971). The Assessment and Analysis of Handedness: the Edinburgh Inventory. *Journal of Neuropsychologia*, *9*, 97-113.

Oviatt, S. Arthur, A., & Cohen, J. (2006). Quiet Interfaces That Help Students Think. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology,* 191-200.

Ozok, A.A., Senson, D., Chakraborty, J., & Norcio, A.F. (2008). A Comparative Study between Tablet and Laptop PCs: User Satisfaction and Preferences. *Journal of Human Computer Interaction*, 329-352.

Park, E., Kim, K. J., & del Pobil, A. P. (2011). Does Length Matter? A Study Examining How Length of Stylus Helps Effective Electronic Documentation. In the *Proceedings of the 4th International Conference on Interaction Sciences*, 185-188.

Perski, H. & Morag, M. (2004). United States Patent No. 6762752 B2, Washington, D.C.: *U.S. Patent and Trademark Office*.

Ramakers, R., Vanacken, D., Luyten, K., Coninx, K., & Schöning, J. (2012). Carpus: a Non-Intrusive User Identification Technique for Interactive Surfaces. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 35-44.

Ramos, G., Boulos, M., & Balakrishnan, R. (2004). Pressure Widgets. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 487-494.

Ramos, G., Cockburn, A., Balakrishnan, R., & Beaudouin-Lafon, M. (2007). Pointing Lenses: Facilitating Stylus Input through Visual-And Motor-Space Magnification. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 757-766.

Ramos, G. & Balakrishnan, R. (2003). Fluid Interaction Techniques for the Control and Annotation of Digital Video. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 105-114.

Ramos, G. & Balakrishnan, R. (2005). Zliding: Fluid Zooming and Sliding for High Precision Parameter Manipulation. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 143-152.

Ramos, G. & Balakrishnan, R. (2007). Pressure Marks. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1375-1384.

Read, J.C. (2006). The Usability of Digital Ink Technologies for Children and Teenagers. *People and Computers XIX-The Bigger Picture*, 19-35.

Ren, X., Yin, J., Zhao, S., & Li, Y. (2007). The Adaptive Hybrid Cursor: a Pressure-Based Target Selection Technique for Pen-Based User Interfaces. In the *Proceedings of INTERACT*, 301-323.

Rosner, D., Oehlberg, L., & Ryokai, K. (2008). Studying Paper Use to Inform the Design of Personal and Portable Technology. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3405-3410.

Sage, S. (2011). Synaptics Series 4 Capacitive Touchscreen Works With Stylus, Through Gloves. Retrieved from http://www.intomobile.com/2011/02/16/synaptics-series-4-clearpad/.

Samsung. (n.d.) How Do I Use Air Gestures? Retrieved July 2013 from http://www.samsung.com/us/support/howtoguide/n0000003/10141/120552.

Sawchuk, A.A., Chew, E., Zimmermann, R., Papadopoulos, C., & Kyriakakis, C. (2003). From Remote Media Immersion to Distributed Immersive Performance. In the *Proceedings of the ACM SIGMM Workshop on Experiential Telepresence*, 110-120.

Schwartz, J., Xiao, R., Mankoff, J., Hudson, S., & Harrison, C. (2014). Probabilistic Palm Rejection Using Spatiotemporal Touch Features and Iterative Classification. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009-2012.

Seow, S. (2008). *Designing and Engineering Time: the Psychology of Time Perception in Software*. Pearson Education, Boston.

Selin, A. (2003). *Pencil Grip: A Descriptive Model and Four Empirical Studies*. Akademi University Press.

Sellen, A. & Harper, R.H.R. (2003). *The Myth of the Paperless Office*. MIT Press Cambridge.

Sharp. (n.d.). Sharp to Release Compact Digital Notepad. Retrieved January 2013 from http://www.cdrinfo.com/sections/news/details.aspx?newsid=35231.

Shu. K. (2013). *Understanding and Rejecting Errant Touches on Multi-Touch Tablets.* (Master's Thesis) Singapore Management University, Bras Basah, Singapore.

Siio, I. and Tsujita, H. (2006). Mobile Interaction Using Paperweight Metaphor. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technologies*, 111-114.

Smartkeitai. (2013). *Sharp. 10.1-Inch IGZO High Sensitivity Tablet Display W. Pen Input*. Retrieved from http://www.youtube.com/watch?v=ofvzmluaibg.

SmudgeGuard. (n.d.). *Eliminates Smudges and Hand Friction for Artists, Left-Handed Children & Adults, Wacom and Tablet PC Users*! Retrieved June 2014 from http://smudgeguard.com.

Sony. (n.d.). *Floating Touch-Developer World Mobile*. Retrieved June 2014 from http://developer.sonymobile.com/knowledge-base/technologies/floating-touch/.

STMicroelectronics. (n.d.). Retrieved June 2014 from http://www.st.com/web/en/catalog/sense_power/fm89/sc1717.

Song, H., Benko, H., Gumbretière, F., Izadi, S., Cao, X., & Hinckley, K. (2011). Grips and Gestures on a Multi-Touch Pen. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1323-1332.

Sun, M., Ren, X., Zhai, S., & Mukai, T. (2012). An Investigation of the Relationship Between Texture and Human Performance In Steering Tasks. In the *Proceedings of the Asia Pacific Conference on Computer Human Interaction Systems*, 1-6.

Sutherland, I. (1963). *Sketchpad, a Man-Machine Graphical Communication System*. (Doctoral Dissertation). Massachusetts Institution of Technology, Cambridge, Massachusetts, Untied States of America.

Suzuki, .Y., Misue, K., & Tanaka, J. (2007). Interface for Digital Notes Using Stylus Motions Made In the Air. In the *Proceedings of Knowledge-Based and Intelligent Information and Engineering Systems*, 104-109.

Suzuki, Y., Misue, K., & Tanaka, J. (2010). Interaction Technique Combining Gripping and Pen Pressures. In the *Proceedings of Knowledge-Based and Intelligent Information and Engineering Systems*, 440-448.

Tchalenko, J. (2007). Eye Movements in Drawing Simple Lines. *Perception*, *36*, 1152-1167.

Tchalenko, J. & Miall, R.C. (2009). Eye-Hand Strategies in Copying Complex Lines. *Cortex*, *45*, 368-376.

Toyoda, N., Yamamoto, R., & Yubuta, T. (2011). Clarification of Transition Conditions for Eye Movement While Generating a Trajectory Using the Upper Limb. In the *Proceedings of the International Symposium on System Integration*, 106-111.

Toyoda, N., Yamamoto, R., & Yubuta, T. (2012). Eye Movement While Generating a Trajectory Using the Upper Limb. *Journal of Energy and Power Engineering*, *6*, 1733-1744.

Taylor, M.M. & Creelman, C.D. (1967). PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America*, *41*, 782-787.

Toy, K.J., Peres, S.C., David, T.Y., Nery, A., & Phillips, R.G. (2012). Examining User Preferences in Interacting With Touchscreen Devices. In the *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1862-1866.

Tu, H., Ren, X., & Zhai, S. (2012). A Comparative Evaluation of Finger and Pen Stroke Gestures. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1287-1296.

Valderrama Bahamóndez, E., Kubitza, T., Henze, N., & Schmidt, A. (2013). Analysis of Children's Handwriting on Touchscreen Phones. In the *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services,* 171-174.

Vogel, D. & Balakrishnan, R. (2010). Direct Pen Interaction With a Conventional Graphical User Interface. *Journal of Human Computer Interaction*, 324-388.

Vogel, D. & Balakrishnan, R. (2010). Occlusion-Aware Interfaces. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 263-272.

Vogel, D. & Baudisch, P. (2007). Shift: a Technique for Operating Pen-Based Interfaces Using Touch. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 657-666.

Vogel, D., Cudmore, M., Casiez, M., Balakrishnan, R., & Keliher, L. (2009). Hand Occlusion With Tablet-Sized Direct Pen Input. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 557-566

Wacom. (n.d.). Inkling. Retrieved June 2014 from http://www.wacom.com/en/ca/creative/inkling.

Wagner, J., Huot, S., & Mackay, W. (2012). Bitouch and Bipad: Designing Bimanual Interaction for Hand-Held Tablets. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2317-2326.

Wald, A. (1947). *Sequential Analysis.* John Wiley & Sons Inc., New York.

Wang, F., Cao, X., Ren, X., & Irani, P. (2009). Detecting and Leveraging Finger Orientation for Interaction with Direct-Touch Surfaces. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 23-32.

Ward, J.R. & Phillips, M.J. (1987). Digitizer Technology: Performance Characteristics and the Effects on the User Interface. *Computer Graphics and Applications*, *7(4),* 31-44.

Wigdor, D. Benko, H., Pella, J., Lombardo, J., & Williams, S. (2011). Rock & Rails: Extending Multi-Touch Interactions With Shape Gestures to Enable Precise Spatial Manipulations. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1581-1590.

Windows for Pen Computing. (n.d.). In *Wikipedia*. Retrieved June 2014 from http://en.wikipedia.org/wiki/windows_for_pen_ computing.

Wintergerst, G., Jagodzinski, R., Hemmert, F., Muller, A., & Joost, G. (2010). Reflective Haptics: Enhancing Stylus-Based Interactions on Touch Screens. In the *Proceedings of the International Conference on Haptics: Generating and Perceiving Tangible Sensations*, 360-366.

Wintergerst, G., Jagodzinski, R., Held, M., Hemmert, F., Muller, A., & Joost, G. (2011). Combining Active and Passive Haptic Feedback for Rich Stylus Interactions. *Mensch & Computer*, *7*.

Wright, J. & Brandt, E. (2001). System-Level MIDI Performance Testing. In the *Proceedings of International Computer Music Conference*, 318-331.

Wu, M.S. (2012). United States Patent No. 20120313883 A1, Washington, D.C*.: U.S. Patent and Trademark Office.*

Wu, M. & Balakrishnan, R. (2003). Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 193-202.

Wu, F. & Luo, S. (2006). Performance Study on Touch-Pens Size in Three Screen Tasks. *Applied Ergonomics*, *37*, 149-158.

Xin. Y., Bi, X., & Ren, X. (2011). Acquiring and Pointing: An Empirical Study of Pen-Tilt-Based Interaction. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 849-858.

Xin. Y., Bi, X., & Ren, X. (2012). Natural User Profiles for the Pen: An Empirical Exploration of Pressure, Tilt, & Azimuth. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 801-804.

Xin, Y., Li, Y., & Ren, X. (2010). A Pen Pressure Division Method Using Personal Distribution. In the *Proceedings of IEEE International Conference on Information and Automation*, 793-798.

Yamanami, T., Funahashi, T., & Senda, T. (1989). United States Patent No. 4878553 A1, Washington, D.C.: *U.S. Patent and Trademark Office.*

Yin, J. & Ren, X. (2008). ZWPS and Pressure Scroll: Two Pressure-Based Techniques in Pen-Based Interfaces. *Information and Media Technologies, 3(1),* 79-90.

Yoon, D., Chen, N., & Gumbretière, F. (2013). TextTearing for Easy Annotaiton of Digital Documents. In the *Proceedings of the ACM Symposium on User Interfaces and Software Technology*, 107-112.

Zabramski, S. (2011). Careless Touch: a Comparative Evaluation of Mouse, Pen, & Touch Input in Shape Tracing Task. In the *Proceedings of the Australian SIGCHI Conference on Human Factors in Computing Systems Conference*, 329-332.

Zeleznik, R., Miller, T., Van Dam, A., Li, C., Tenneson, D., Maloney, C., & Laviola, J.J. (2012). Applications and Issues in Pen-Centric Computing. *Multimedia, 15(4),* 14-21.

Zhang, H., Yang, X.D., End, B., Liang, H.N., Boulanger, P., & Irani, P. (2012). See Me, See You: A Lightweight Method for Discriminating User Touches on Tabletop Displays. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2327-2336.

Zitnick, L. (2013). Handwriting Beautification Using Token Means. *ACM Transactions on Graphics*, *32(4)*, 53.