The Pen Is Mightier: Understanding Stylus Behaviour While Inking on Tablets

Michelle Annett *
University of Alberta
Microsoft Research

Fraser Anderson †
University of Alberta

Walter F. Bischof §
University of Alberta

Anoop Gupta ¥
Microsoft Research

Abstract

Although pens and paper are pervasive in the analog world, their digital counterparts, styli and tablets, have yet to achieve the same adoption and frequency of use. To date, little research has identified why inking experiences differ so greatly between analog and digital media or quantified the varied experiences that exist with stylus-enabled tablets. By observing quantitative and behavioural data in addition to querying preferential opinions, the experimentation reaffirmed the significance of accuracy, latency, and unintended touch, whilst uncovering the importance of friction, aesthetics, and stroke beautification to users. The observed participant behaviour and recommended tangible goals should enhance the development and evaluation of future systems.

Keywords: Tablet; Stylus; User Interaction; Accuracy; Latency; Stylus Design; Palm Rejection; Unintended Touch.

Index Terms: H.5.2. User Interfaces: Input devices and strategies

1 Introduction

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 direct manipulation, tablets should be ideal devices for productivity-based activities such as note taking, sketching, or annotation. Work by Muller et al., however, demonstrated that tablets are championed for content consumption activities such as gaming, web browsing, social networking, and email instead of inking-based activities [21].

When coupled with touch, a stylus enabled tablet should afford efficient bimanual interaction, supporting the transfer of behaviours and interaction techniques commonly found with traditional pen and paper [1, 9, 10, 11, 12]. By harnessing fine motor control, styli offer increased precision and accuracy compared to the fingers [4, 13, 19, 31], allowing users to diagram, work out equations, sketch, create calligraphy, annotate, or sign their name in a manner more natural and fluid than with a mouse or keyboard [36]. Even with these benefits, consumer and educational usage of tablets for these tasks unfortunately remains low.

Some manufacturers have acknowledged the potential usefulness of pens, including styli with many newer models. Such styli typically come in two varieties, *active* and *passive*. *Passive styli* use capacitive sensing to detect touch and do not provide support for pressure, mode switching, etc. They are often after-market peripherals that mimic the properties of the finger on capacitive panels so they cannot be differentiated from touch. Although passive styli are often afterthoughts and introduce many usability problems, they are increasingly being used with current and legacy devices. *Active styli*, however, require special digitizing hardware to detect the stylus independently from touch. Such styli are typically more precise, respond to pressure, and have barrel buttons for mode switching, making them suitable for most inking tasks. 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.

Although pen computing has a long history, most work has assessed specific issues with active styli such as bi-manual interaction [1, 10, 11, 12, 20] or stylus features (e.g., pressure, tilt, azimuth [2, 27]). However, exploratory work by Vogel and Balakrishnan 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 [33]. Within the literature, there is unfortunately little empirical evidence regarding the problems users face while *inking* (e.g., drawing and writing) Similarly, the behavioural adaptations necessary to accomplish routine inking tasks, issues with passive styli, or how to best identify or evaluate inking issues remain unknown.

To understand the problems experienced by tablet users, we observed how participants used traditional pen and paper, as well as 'best' and 'worst' case digital devices, while sketching and note taking. While the digital and analog experiences are not believed to be identical, paper provides an excellent gold standard and baseline of 'frustration free' inking experiences in terms of comfort and efficiency to compare to. By observing behaviour generated from real world activities, inspecting content created by participants, and analysing questionnaires, the greatest sources of frustration while inking with styli-enabled systems were uncovered.

Specifically, this paper contributes:

- An analysis of behavioural and performance differences that occurred while inking with digital and analog media
- The description and classification of behavioural adaptations, hand movements, and grips unique to tablets while writing and sketching
- Identification and prioritization of outstanding issues prohibiting satisfying stylus experiences
- A set of tasks and quantitative measures to assess a device's suitability for inking tasks

2 Related Work

Throughout the history of pen computing, much work has sought to understand 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.

A number of projects explored usage patterns and behaviour with tablets. Muller, Gove, and Webb conducted a multi-method exploration of tablet usage and found that activities such as checking email, playing games, and social networking were most common [21]. Wagner and colleagues 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

* email: mkannett@ualberta.ca
† email: frasera@ualberta.ca
§ email: wfb@ualberta.ca
¥ email: anoop@microsoft.com

fingers and palm in consistent locations on or around the tablet [34]. Toy et al. 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 [30]. These investigations have underscored the lack of support for styli and inking tasks in the consumer tablet experience. The present work analyses why such a disparity occurs and prioritizes the of the stylus experience that are most important to users.

Many researchers have focused on identifying those tasks the stylus is most beneficial for. Device and task interactions have been largely confirmed, with the stylus being optimal for compound tasks, crossing tasks, radial steering, selection, stroke-based gestures, and shape tracing tasks [4, 8, 13, 18, 31, 35]. The stylus was also found to produce more legible notes than a finger on mobile phones [32]. Work by Briggs and colleagues and Ozok et al. investigated preferences for the different tasks and activities supported by a stylus [3, 26]. 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. Although styli have yet to be widely adopted, such work illustrates the superiority of styli for many tasks. In contrast to this work, we focus on the natural writing and sketching behaviours that occur on tablets, without the use of excessive stroke beautification or handwriting recognition.

To enhance the stylus experience, some researchers have focused on examining pen and paper usage to inform the design and features of digital devices. For example, Hinckley et al. observed participants' preferred and non-preferred hand usage of a paper notebook, pens, glue, scissors, and paper content while creating a storyboard [11, 12] to inform the interaction techniques used in Manual Deskterity. Work by Fitzmaurice and colleagues compared how often participants rotated paper, a tethered tablet, and a 6DOF tablet while completing two handwriting and three drawing tasks, using the results to inform the design of rotating user interfaces [6]. Rosner and colleagues surveyed existing literature, highlighting the importance of the physicality of paper, the folding and dog-earing of pages, the use of dust jackets, and the placement of tabs in notebooks [28]. Lim explored the differences in thinking processes and cognitive behaviour on architects' ability to compose drawings using styli versus pen and paper [16]. More time was spent inspecting and exploring the digital drawings than paper ones, calling for an increased usage of stylus-based systems within architecture. Oviatt, Arthur, and Cohen explored how cognitive load and performance was affected by using digital devices as well as pen and paper devices, and found that the more interfaces deviate from paper, the greater the cognitive load [25]. Similar to this work, we observed natural pen and paper behaviour, but in contrast, we explore behaviour and adaptations on digital devices to inform future stylus experiences. There is a need for the identification of behavioural changes that occur when one uses a tablet instead of pen and paper and an exploration of the underlying causes for these behaviours. The present work explicitly focuses on understanding the hardware and software features that should be improved to provide a satisfying digital inking experience.

3 Experimental Methodology

A user study was conducted to identify the behavioural, performance, and preferential differences brought about while inking with analog and digital media. The exploration was not intended to demonstrate that digital styli provide a better inking experience than pens (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'. The inclusion of passive styli provided verification that the tasks and measures accurately reflected the tablet experience and allowed for the exploration of the full range of behavioural adaptations made with current commercial tablets.

3.1 Participants

Thirty participants (10 female) were recruited for the study (M = 39, SD = 10 years). The Edinburgh Handedness Inventory [24] 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.

3.2 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 the capacitive devices are not designed for stylus input, the iPad was included as it enabled an evaluation of the 'worst case' of stylus experiences and it has an expanding ecosystem of third party passive styli. The active stylus device, the Slate, was specifically designed with stylus support in mind and it is expected to provide a better experience for the user. Other form-factors such as the Wacom Cintiq, Intuos, and Bamboo, were intentionally excluded to examine form-factors commonly used by everyday users, not niche equipment used by experts and professionals.

Three Casio ZR100 cameras were setup to capture participant behaviour (Figure 1). One camera was located above the participant and recorded the entire interaction area. As both left and right-handed participants participated, one of two side cameras were used to capture the vertical movement of the styli and hands.

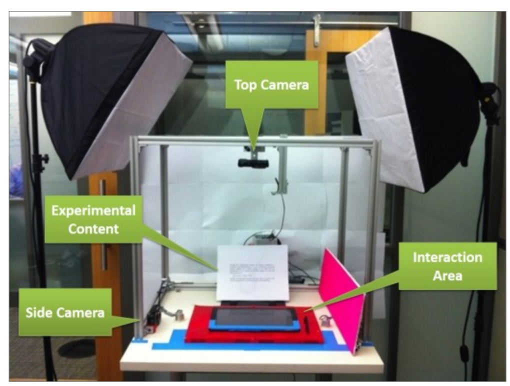

Figure 1: The experimental setup with the locations of the cameras, experimental content, and interaction areas highlighted.

To enhance ecological validity and repeatability, popular, freely available inking applications on the iPad (*Noteability*) and Slate (*PDF Annotator*) instead of custom or professional programs. For inking, participants were provided with a Wacom Bamboo Solo passive stylus with a nib thickness of 7 mm to ink and a Samsung ATIV Tablet S-Pen active stylus with a nib thickness of 2 mm. Uniball ONYX Fine pens with a nib thickness of 0.7 mm were provided for the paper-based conditions. Across all three media, the ink line thickness was approximately 0.7 millimetres 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.

3.3 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 containing a mathematical equation (Figure 2). An equation was included to capture scenarios where non-traditional, unfamiliar symbols are used. Participants completed a transcription task rather than generating their own content to ensure behaviour was comparable between participants and across media and to impose divided attention, which is common during real-world writing.

The product of a sequence is derived using the letter Π from the Greek alphabet (similar to the ∑ use in summation). A product is defined as:

$$\prod_{i=m}^{n} x_i = x_m \cdot x_{m+1} \cdot x_{m+2} \cdot \cdots \cdot x_{n-1} \cdot x_n$$

The factors of the product are obtained by substituting successive integers for the index of multiplication, starting from the lower and incrementing to the upper bound.

Mathematical notation uses ∑, an enlarged form of the Greek letter Sigma, to represent the summation of many similar terms. For example, in

$$\sum_{i=m}^{n} x_i = x_m + x_{m+1} + x_{m+2} + \cdots + x_{n-1} + x_n$$

I represents the index of summation, x_i is an indexed variable representing each successive term in the series, and m and n are the lower and upper bounds of summation.

If f is a continuous real-valued function defined on a closed interval [a, b], then once an antiderivative F of f is known, the definite integral of f over that interval is given by:

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

Integrals and derivatives are the basic tools of calculus. They have numerous applications in computing science, mathematics, and engineering.

Figure 2: The stimuli used for the experiment: (left) The sample paragraphs containing alphanumeric and mathematical content and (right) the organic figures used for the sketching tasks.

With sketching, participants copied an organic figure, capturing as many details as possible in five minutes (Figure 2). Each figure contained strokes of varying lengths and directions and had explicit shading regions that required straight and curved lines. In pilot testing, photographs and real objects lead to huge variations in behaviour, with 30 second to 5 minute sketches. Sketching a unique shape with an explicit outline and shading ensured the task and movements were consistent between participants.

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

3.4 Measures and Data Analysis

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

3.4.1 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, etc.) easily identifiable behaviours that had a direct impact on comfort were analysed: grip and movement style. For grip, stylus grips and hand postures previously identified in the literature were consulted [15, 29]. To quantify the hand accommodations, the video data was manually analysed by one of the authors to find the 'nearest match' for each participant and each 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 hand 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 it remained within the same grip category during the task. To simplify the presentation of results, grip and hand-movement are provided at a per-task level.

3.4.2 Writing Size

The most appropriate measure found to quantify the effects of each device's characteristics on user behaviour during a pilot study was writing size. To compute writing size, the average height from the baseline to x-height [7] of each line was computed by averaging the height at three different points along each line (not including the larger sigma or pi characters). The writing size of the line containing the equation, *equation line*, was computed separately from the other lines, *text lines,* resulting in two measures of writing size. Writing size measures are presented in points to be consistent with common typographic conventions; twelve points are equivalent to 4.2 millimeters, as measured on-screen.

3.4.3 User Preferences While Inking

As perception and opinions towards devices are important, ratings regarding the appropriateness of using each medium were collected for both 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.

4 Results and Discussion

For each measure, the observed behaviours are detailed and followed by the statistical analysis and interpretation 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 favoured. We do not claim to provide definitive claims regarding universal properties of active or passive devices, as there are many differences between the media aside from digitizing technology.

4.1 Hand Accommodations

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

4.1.1 Grip

Prior work on grips has identified a variety of grips common during content creation tasks (e.g., writing, drawing, painting, tool use). Levy and Reid identified one grip, the *natural* grip, and described an 'inversion' behaviour that some subjects, especially those that were left handed, performed [15]. Song et al. described five grips, i.e., *tripod*, *relaxed tripod*, *sketch*, *tuck*, and *wrap*, observed during a pilot exploration of four participants with Wacom devices [29]. These five grips are largely task dependent (e.g., the wrap grip used only for painting, the sketch grip to shade with a pencil lead) and two were similar to the natural grip previously identified by Levy and Reid (i.e., tripod and relaxed tripod).

Of these six grips identified in the literature, only one was observed during the experiment: the *natural* grip [15, 29]. Two other grips, the *knuckle* and *crab* grips (depicted in Figure 3, usage in Figure 4) were identified as 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 an *other* category.

Figure 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).

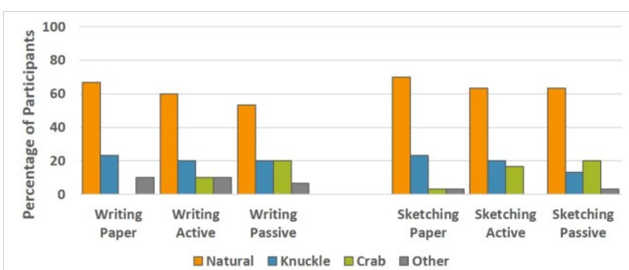

Figure 4: The grips displayed by participants while writing and sketching, presented by task and medium. Note the prevalence of the 'crab' grip while using digital devices.

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 were found influence 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, the type of device ($p < 0.05$), task ($p < .001$) and handedness of the participant ($p < .001$) 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 ($p < .001$).

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 was used almost exclusively with digital devices, 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 the importance of implementing palm rejection. Unlike active stylus systems, passive stylus systems cannot predict where the stylus will be, so palm rejection becomes much more difficult and encourages 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 commonly adopted by left-handed writers [15], those that exhibited the knuckle grip also rotated the digital media 20 - 40 degrees. Identification of the knuckle grip and the accompanying medium rotation is important for designers of unintended touch solutions to note, as the palm would likely produce a very different pattern of sensor activation that with the natural grip.

4.1.2 Hand Movement Style

Participants exhibited one of three categories of hand movement patterns: *floating*, *planting*, or *dragging* (as depicted in Figure 5). The most prominent behaviour, *floating*, was exhibited by 51% of the 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.

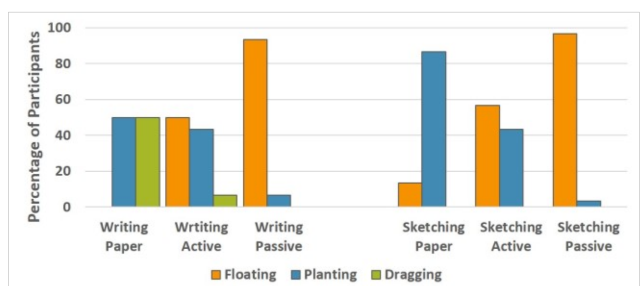

Figure 5: The patterns of hand movements used by participants while writing and sketching. Note the lack of floating with paper and the lack of dragging with digital devices.

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 media ($p < .001$), task ($p < .001$), and handedness ($p < .05$) 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 system ($p < .001$) and active 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 surface of paper 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, on the passive system almost all participants modified their behaviour. The difference in the floating movements on the passive versus active systems suggests that the active system's identification and rejection of unintentional touch events was unacceptable for most participants 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 participants are able to transfer their normal writing behaviours onto digital devices. It is possible that those who lifted their palm when using the active stylus and Slate were pre-conditioned to lift their palm by prior experiences with passive styli or other touchscreen devices.

4.2 Writing Size

To evaluate the character size used while writing (Figure 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: $F(1,27) = 0.074$, $p = .788$; Equation Line: $F(1,27) = 0.207$, $p = .652$), so the handedness factor was collapsed and another ANOVA was performed without this factor. This second ANOVA determined that the device used influenced the writing size (Text lines: $F(2,50) = 9.958$, $p < .001$; Equation Line: $F(1.6,39.5) = 12.840$, $p < .001$). Post-hoc pairwise comparisons determined that participants wrote smallest on paper, slightly larger on the active device, and largest with the passive device (Table 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.

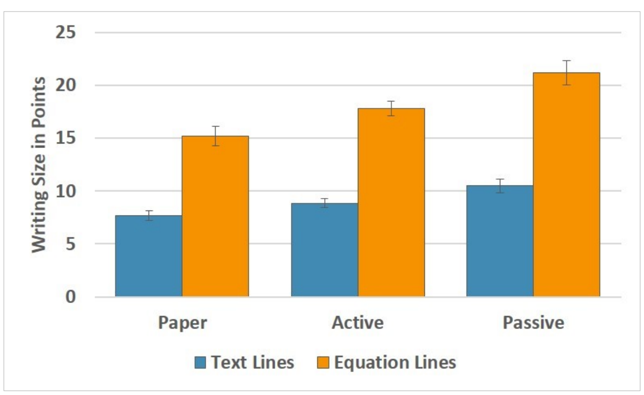

Figure 6: Writing size used on each device for the text lines and equation lines. Note the increase from paper, to active and passive. Error bars represent the standard error of the mean.

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 by the hover state about the presence and location of the nib 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."

Table 1: 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 *

4.3 User Preferences While Inking

No significant differences were found between the handedness groups with respect to the Likert-scale ratings (Figure 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), with paper being the most preferred, followed by the active and then passive system.

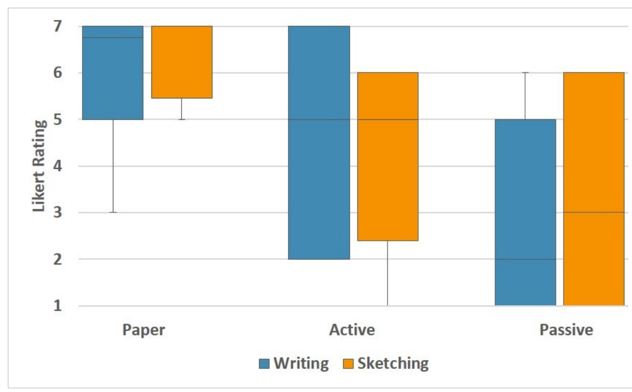

Figure 7: Participant median responses to "I feel that the {Paper, Slate, iPad} was a good medium to complete the {sketching, writing} task with". Note the decline in ratings from paper, to active and passive.

Table 2: 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 *

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 and this was reflected in participants' ratings. As active stylus systems are optimized for inking, the active system was rated higher than the passive system (i.e., the median response was 'Slightly Agree').

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 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 device, and lastly the passive device. While this is expected, this supports the validity of the tests and measures used.

5 FEATURES INFLUENCING THE STYLUS EXPERIENCE

Our experiments uncovered many elements that influence behaviour, performance, and preferences for digital versus analog media while inking. From the observations and participant comments, many features impacting usability emerged, with participants identifying five that present substantial, pressing issues. 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 that each had. The identification these primary and secondary features, as well as differences between the tablet devices, should decrease the need for adaptations in the future and enhance the design tablet hardware and software.

5.1 Primary Features

Stylus accuracy, device latency, and unintended touch resulted in the greatest effect on participant behaviour and were the most prominent features identified as problematic by participants.

5.1.1 Stylus Accuracy

A recurring theme that emerged was frustration due to inaccuracy of the stylus. Many participants were vocal about inaccuracy, as it forced them to alter their writing size. Participants had more difficulty forming and terminating letters with the passive system than the active system and paper (Figure 8). The hover information provided by the active system, 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 8).

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 calibration, interpolation, parallax, and sensor density influence device accuracy, the physical design of the stylus appears to 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 imperative that increased accuracy, intuitive navigation methods (not accidently activated if users rest their hand or fingers on the screen), and intelligent widgets or canvases capable of reflowing or reformatting content as necessary, become integrated within stylus-supported applications.

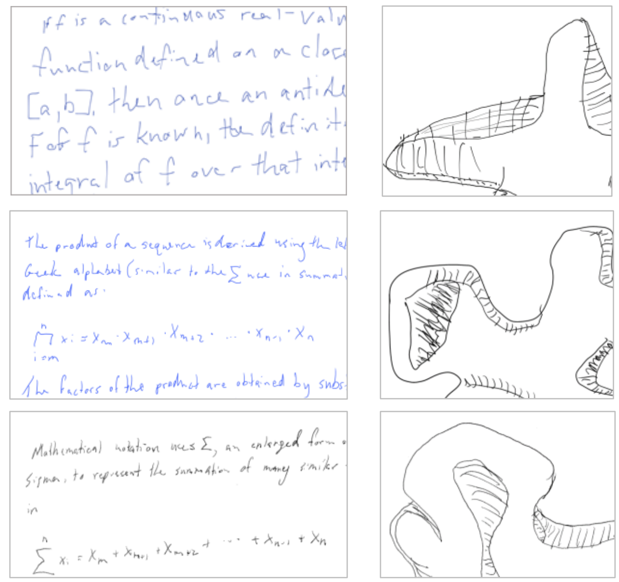

Figure 8: Inking content created using the passive stylus (top), active stylus (middle), and pen and paper (bottom) from the same participant. Note the inaccuracy of the lines on both digital devices, as well as the limited use of curved lines with the passive device. The images were cropped to show detail, and the paper image was scanned, resulting in the perceived loss of quality.

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 detect hover, nib-size, cursor calibration, the texture of the screen, the material composition of the stylus, and the responsiveness of the device. Although third-party styli such as Project Mighty, Pencil, and Jot Pro have begun to use auxiliary input channels to improve some of these factors for passive systems, there remains much work for designers of both passive and active systems. Further experimentation is still needed to tease apart these factors to better understand the inaccuracies users can tolerate as well as how to minimize inaccuracy overall.

5.1.2 Device Latency

One hundred milliseconds has long been regarded as the minimum latency necessary for satisfying direct-interaction [[19]]. Although the iPad and Slate had end-to-end latencies below this threshold (80 and 65 milliseconds respectively, computed using the method in [23]), 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 resulting in participants writing 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 [14]. 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. Although we did not explicitly focus on latency, further work must investigate the relationship between latency and accuracy to determine acceptable standards and benchmarks for high accuracy tasks such as inking. Until latency is decreased across the entire pipeline, low-fidelity, high-frequency feedback should be provided. Instead of taking the time to render a smoothed, high-quality line, for example, initially rendering a quick, crude stroke and later replacing it with a smoothed line when more processing is available may be fruitful to consider. 3D 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.

5.1.3 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". Participants did not experience much fatigue, as they were only inking for 5 minutes. Inking for longer periods would have likely exacerbated fatigue and 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 (interleaved interaction). Future work should focus on palm rejection and improving the identification of unintended touch. In most applications and systems today, users do not have the opportunity to provide personalized information about their handedness or grips. The different grips that are used, especially those unique to left-handers, will produce different imprints on a sensor array and could be harnessed for palm rejection. Until unintended touch is solved, either via hardware improvements or software solutions, designers should provide mechanisms by which users can enable or disable touch input if desired and also provide feedback to alert users of any stray marks that are rendered.

5.2 Secondary Features

Surface and nib texture, stylus and screen aesthetics, and stroke beautification were also noted as being important to the stylus experience by participants, albeit to a lesser extent than the primary features.

5.2.1 Texture and 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 mismatch of the friction between the hand and surface. Many participants felt that "there was not enough friction between the pen and screen to feel natural". This mismatch between skin and surface friction was also reflected in comments such as, "my hand jerked across the screen as I moved it" and 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 materials the stylus nib is composed of should be optimized to evoke familiar feedback patterns for the user and encourage natural movement instead of hindering it.

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 writing 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". None of the participants mentioned the increased thickness or weight of the digital devices compared to paper. This is likely because that 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 colours (i.e., black or grey) and the choice of nibs is limited (expect with active Wacom 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 can invoke a stronger connection to one's work, which is a natural strength 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.

5.2.2 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 [5, 17, 37], the current beautification methods employed for inking obviously did not meet participant's expectations.

These comments highlight the value of pressure sensitivity and 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, and azimuth in ink rendering, not only within the context of pressure-based widgets or the levels of tilt, azimuth, or pressure discernible [2, 27]. Such improvements will uphold beliefs that tablets can provide experiences similar, if not more appropriate and engaging for productivity-based tasks.

6 Conclusion

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

Our work 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 are in need of future work. 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. Latency, surface texture, unintended touch (palm rejection), and input resolution have not seen significant advancements in recent years. The tablet and stylus has great potential to become 'go-to' devices for inking and productivity-based activities, but many improvements are needed before tablets and styli become commonplace in everyday settings.

References

[1] R. Balakrishnan and K. Hinckley. The Role of Kinesthetic Reference Frames in Two-Handed Input Performance. In *Proc. of UIST*, 171-178, 1999.

[2] X. Bi, T. Moscovich, G. Ramos, R. Balakrishnan, and K. Hinckley. An Exploration of Pen Rolling For Pen-Based Interaction. *In Proc. of UIST*, 191-200, 2008.

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

[4] A. Cockburn, D. Ahlstrom, and C. Gutwin, Understanding Performance in Touch Selections: Tap, Drag, and Radial Pointing Drag with Finger, Stylus, and Mouse. *Journal of Human-Computer Studies*, 70: 218-233, 2008.

[5] J.D. Fekete, É. Bizouarn, É. Cournarie, T. Galas, and F. Taillefer. Tictactoon: A Paperless System for Professional 2D Animation. In *Proc. of SIGGRAPH*, 79-95, 1995.

[6] G. Fitzmaurice, R. Balakrishnan, G. Kurtenbach, and B. Buxton. An Exploration into Supporting Artwork Orientation in the User Interface. In *Proc. of CHI*, 167-174, 1999.

[7] Font Shop Education. *http://www.fontshop.com/education/pdf/typeface_anatomy.pdf.* Accessed August 2013.

[8] C. Forlines and R. Balakrishnan. Evaluating Tactile Feedback and Direct vs Indirect Stylus Input in Pointing and Crossing Tasks. In *Proc. of CHI*, 1563-1572, 2008.

[9] Y. Guiard. Asymmetric Transfer of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. *Journal of Motor Behaviour*, 19: 486-517, 1987.

[10] K. Hinckley. Input Technologies and Techniques. *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, Taylor and Francis Group, LLC, New York, 2002.

[11] K. Hinckley, K. Yatani, M. Pahud, N. Coddington, J. Rodenhouse, A. Wilson, H. Benko, and B. Buxton. Pen + Touch = New Tools. In *Proc. of UIST*, 27-36, 2010.

[12] K. Hinckley, K. Yatani, M. Pahud, N. Coddington, J. Rodenhouse, A. Wilson, H. Benko, and B. Buxton. Manual Deskterity: An Exploration of Simultaneous Pen + Touch Direct Input. In *Extended Abstracts of CHI*, 2793-2802, 2010.

[13] A. Holzinger, M. Holler, M. Schedlbauer, and B. Urlesberger. An Investigation of Finger versus Stylus Input in Medical Scenarios. In *Proc. of ITI*, 433-438, 2008.

[14] R. Jota, A. Ng, P. Dietz, and D. Wigdor. How Fast Is Fast Enough? A Study of the Effects of Latency in Direct-touch Pointing Tasks. In *Proc. of CHI,* 2291-2300, 2013.

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

[16] C. Lim. An Insight into the Freedom of Using a Pen: Pen-Based System and Pen-And-Paper. In *Proc. of 6th Asian Design International Conference*, 2003.

[17] Y. Lu, F. Yu, A. Finkelstein, and S. DiVerdi. Helping Hand: Example-based Stroke Stylization. *ACM Transactions on Graphics*, 31(4), 1-10, 2012.

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

[19] I.S. Mackenzie, A. Sellen, and B. Buxton. A Comparison of Input Devices in Elemental Pointing and Dragging Tasks. In *Proc. of CHI*, 161-166, 1991.

[20] F. Matulic and M. Norrie. Empirical Evaluation of Uni- And Bimodal Pen and Touch Interaction Properties on Digital Tabletops. In *Proc. of ITS,* 143-52, 2012.

[21] R.B. Miller. Response Time in Man-Computer Conversational Transactions. Fall Joint Computer Conference, 267-277, 1968.

[22] H. Muller, J.L. Gove, and J.S. Webb. Understanding Tablet Use: A Multi-Method Exploration. In *Proc. of MobileHCI*, 1-10, 2012.

[23] A. Ng, J. Lepinski, D. Wigdor, S. Sanders, and P. Dietz. Designing for Low-Latency Direct-Touch Input. In *Proc. of UIST*, 453-462, 2012.

[24] R.C. Oldfield. The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Journal of Neuropsychologia*, 9: 97-113, 1971.

[25] S. Oviatt, A. Arthur, and J. Cohen. Quiet Interfaces that Help Students Think. In *Proc. Of UIST,* 191-200, 2006.

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

[27] G. Ramos and R. Balakrishnan. Pressure Marks. In *Proc. of CHI*, 1375-1384, 2007.

[28] D. Rosner, L. Oehlberg, and K. Ryokai. Studying Paper Use to Inform the Design of Personal and Portable Technology. In *Proc. of CHI,* 3405-3410, 2008.

[29] H. Song, H. Benko, F. Guimbretière, S. Izadi, X. Cao, and K. Hinckley. Grips and Gestures on A Multi-Touch Pen. In *Proc. of CHI*, 1323-1332, 2011.

[30] K.J. Toy, S.C. Peres, T.Y. David, A. Nery, and R.G. Phillips. Examining User Preferences in Interacting with Touchscreen Devices. In *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, 1862-1866, 2012.

[31] H. Tu, X. Ren, and S. Zhai. A Comparative Evaluation of Finger and Pen Stroke Gestures. In *Proc. of CHI,* 1287-1296, 2012.

[32] E. del Carmen Valderrama Bahamóndez, T. Kubitza, N. Henze, and A. Schmidt. Analysis of Children's Handwriting on Touchscreen Phones. In *Proc. of MobileHCI,* 171-174, 2013.

[33] D. Vogel and R. Balakrishnan. Direct Pen Interaction with a Conventional Graphical User Interface. *Journal of Human Computer Interaction*, 324-388, 2010.

[34] J. Wagner, S. Huot, and W. Mackay. BiTouch and BiPad: Designing Bimanual Interaction for Hand-held Tablets. In *Proc. of CHI*, 2317-2326, 2012.

[35] S. Zabramski. Careless Touch: A Comparative Evaluation Of Mouse, Pen, And Touch Input In Shape Tracing Task. In *Proc. of OzCHI*, 329-332, 2011.

[36] R. Zeleznik, T. Miller, A. Van Dam, C. Li, D. Tenneson, C. Maloney, and J.J. LaViola. Applications and Issues in Pen-Centric Computing. *Multimedia,* 15(4):14-21, 2008.

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