How Low Should We Go? Understanding the Perception of Latency While Inking

Michelle Annett [*]
University of Alberta
Microsoft Research

Albert Ng [†]
Applied Sciences Group, Microsoft

Paul Dietz [‡]
Microsoft Research

Walter F. Bischof [§]
University of Alberta

Anoop Gupta[¥]
Microsoft Research

ABSTRACT

Recent advances in hardware have enabled researchers to study the perception of latency. Thus far, latency research has utilized simple touch and stylus-based tasks that do not represent inking activities found in the real world. In this work, we report on two studies that utilized writing and sketching tasks to understand the limits of human perception. Our studies revealed that latency perception while inking is worse (~50 milliseconds) than perception while performing non-inking tasks reported previously (~2-7 milliseconds). We also determined 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 prior and current work has informed the Latency Perception Model, a framework upon which latency knowledge and the underlying mechanisms of perception can be understood and further explored.

Keywords: Latency; delay; responsiveness; stylus; pen; indirect interaction; direct interaction; perception; psychophysics; just-noticeable difference; latency perception model.

Index Terms: B.4.2 Input / Output Devices, H.5.2 User Interfaces: Input devices and strategies, H.1.2. User/Machine Systems: Human factors.

1 INTRODUCTION

While interacting with devices, whether inking, gesturing, or selecting, it is essential to receive feedback about our actions. Users readily notice whenever feedback is unavailable or delayed, adapting their behaviour or interaction styles [26]. On digital devices today, there is typically a 60 to 120 millisecond delay from when the stylus touches the screen until digital ink appears [23]. Caused by both hardware and software factors, such delays decrease user performance [3, 13, 32]. Unlike digital devices, inking with pen and paper incurs zero delay, with the ink flowing from the nib onto the paper, providing instantaneous feedback.

Until recently [3], it was assumed that the delay or latency adequate for direct interaction was on the order of 100 milliseconds [20]. Due to technological advances with high performance hardware, it is now possible to display and examine the influence and perceptibility of much smaller delays, e.g., 1 millisecond, to users [22, 23, 24]. Such hardware has been used to determine the minimum latency required for direct-touch interaction. For example, while tapping with the finger, participants could not distinguish 1 versus 63 milliseconds of delay [13]. While performing a simple box moving task, users could not discriminate 1 versus 6 milliseconds when using the finger [23], but could distinguish 1 versus 2 milliseconds while using a stylus [24].

Although we have learned a great deal about the capabilities of human perception, little is known about latency when increased task complexity and attentional demands are present. Unlike dragging or tapping a target, writing is a cognitively and visually demanding task requiring focus on character size and formation, attention to inter-word spacing, and the ability to ignore cognitive and environmental distractions. Little is known however, about the perceptual processes underlying latency perception or the influence of such conditions on the perception of latency.

In recent work, while performing a stylus-based task that requiring oscillating, scribbling movements, participants were unable to discriminate 7 from 40 milliseconds of latency [24]. Although there is an ecosystem-wide push to make systems faster, such work suggests that these efforts may not be needed. If users perceive two different latencies as being equal, while performing tasks more complex and demanding than tapping, moving a box, or scribbling, it may not be necessary to allocate resources to achieve the latency levels recommended previously. It may be beneficial to reallocate CPU or GPU cycles to improve stroke rendering, integrate pressure or tilt information from the stylus, or improve the recognition of unintended touch. If users are unable to perceive the difference in speed, but readily perceive more realistic looking strokes, delaying ink by 10 or 15 milliseconds may be acceptable if the user experience is improved in other ways.

To understand latency within the context of real-world activities and work towards an understanding of the influencing factors, two experiments were conducted. The aim was not to determine how latency influences performance or develop methods to reduce hardware or software delays, but rather to understand the human perception during real world activities. The first experiment determined the minimal perceivable latency while drawing and writing. Perceived latencies were found to be higher than those previously found with simple dragging and tapping tasks. The experiment also uncovered the importance of task demands and the strategies employed to discriminate latency. These findings motivated the second experiment, which manipulated the location and presence of visual cues and feedback. 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. The results from both 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.

2 RELATED WORK

There has been a variety of work focused on the detection and understanding of latency. Researchers within computer music have strived to determine the ideal latency for musical composition. Early work by Freed et al. and Wright et al. suggested that music controllers should have less than 10 milliseconds latency, as it is at this point that piano players notice delays [9, 34]. Many others have

---

[*] email: mkannett@ualberta.ca
[†] email: v-albeng@microsoft.com
[‡] email: paul.dietz@microsoft.com
[§] email: wfb@ualberta.ca
[¥] email: anoop@microsoft.com

Graphics Interface Conference 2014
7-9 May, Montreal, Quebec, Canada

recommended much higher latencies, depending on the type of instrument, genre of music, and the presence of tactile feedback. Maki-Patola and Hamalainen determined that delays between 2 and 30 milliseconds were sufficient while playing a Theremin without tactile feedback [17]. Adelstein et al. found that delays of 24 milliseconds were tolerable while tapping a brick with a hammer [1]. Dahl and Bresin recommended latencies of 55 milliseconds for percussion instruments [6]. While playing collaborative music via network connection, delays of 100 milliseconds were acceptable while playing piano, but only 20 milliseconds while playing an accordion [25]. Using this work as a guide, Montag and colleagues built a low-cost multi-touch tabletop capable of providing low latency audio and haptic feedback to users for music applications [21]. The system focused on improving 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 [16].

Touch and stylus systems have also been used to investigate the perception of latency. Using off-the-shelf touchscreen devices, Anderson, Doherty, and Ganapthy artificially inserted latency into everyday tasks such as reading and web or photo browsing to determine that users tolerated 80 to 580 milliseconds of delay [4]. Work by Ng et al. focused not on tolerance, but on the minimum latency perceivable [23]. By integrating a high-speed DLP projector with a custom high-speed touch sensor they were capable of displaying touch latencies as low as one millisecond. Ng and colleagues determined that users were able to distinguish between one and six milliseconds of latency while dragging a box. Extensions of this work by Jota et al. examined latency while pointing and during the ‘land-on’ segment of a dragging task [13]. Performance degraded as latency increased during a pointing task and unlike Ng et al.’s box dragging task, participants could not distinguish between 1 and 64 milliseconds during the ‘land-on’ event. Although Ng and Jota’s work probed the boundaries of touch-based latency perception, little information regarding the mechanisms and processes underlying latency perception was provided, nor were stylus-based interactions considered.

While working with a light-pen system, Miller hypothesized that users could tolerate delays on the order of 100 milliseconds when making slow, thoughtful strokes [20]. Although this was the first exploration of latency while using a ‘stylus’, little justification for this estimate is available. Due to technological limitations, stylus latency has only recently been explored again. Henzen et al. developed an electronic ink display that exhibited zero parallax and minimum latency of 40 milliseconds, but did not leave the proof of concept stage [11, 12]. Unlike Henzen and colleagues’ system, our prototype High Performance Stylus System (HPSS) system exhibited zero parallax, and is capable of a one millisecond inking delay while rendering simple shapes and a seven millisecond delay while inking. In previous work with the HPSS, we examined latency perception while participants performed Ng et al.’s box dragging task [23] as well as a scribbling task [24]. We found that participants were able to discriminate between 1 and 2 milliseconds of latency while dragging and 7 and 40 milliseconds of delay while scribbling. As latency perception thus appears to be largely task-dependent and depend upon a multitude of factors, the current study extends our previous work with the HPSS to focus on real-world tasks that have added cognitive load, different loci of attention, and the presence or absence of visual feedback of the hand.

3 JUST NOTICEABLE DIFFERENCE

As the goal of this work was to determine the minimum latency perceivable, it was thus appropriate to use a just-noticeable difference (JND) methodology. With a JND methodology, participants are presented with two stimulus levels and are forced to make a judgment regarding which alternative was brighter, quieter, or in our case, faster. After repeated presentations of various stimuli, one is able to derive the minimum threshold, or *just-noticeable difference* (JND), 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 was constant across trials, we assume that participants would be unable to distinguish between any latencies lower than the threshold.

During our experiments, two latencies 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 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 [13, 23]. Given that our 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 [27] 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 sequential likelihood-ratio test [33] 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, in our case 1 millisecond, the experiment concluded (aka the Minimal Overshoot and Undershoot Sequential Estimation technique [18]). This ensured that participants experienced the smallest possible difference between latencies. A maximum latency upper bound was placed on the probe (i.e., 105 milliseconds) to prevent participants from experiencing higher, unreasonable levels of latency. 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 correctly identified the baseline latency on 75% of trials. As the onscreen digital ink required filtering and smoothing via a moving average window, the minimum latency while inking was seven milliseconds. The baseline latency was thus set at seven milliseconds. 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. Motivated by prior work [13, 23], the initial testing latency was set to 55 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.

4 EXPERIMENT 1: PERCEIVED LATENCY WHILE INKING

To understand latency perception during scenarios that require increased cognitive and attentional demands compared to the tasks used in prior work [24], the first experiment determined the minimum latency perceivable while participants drew simple lines, wrote a word, and sketched a simple shape.

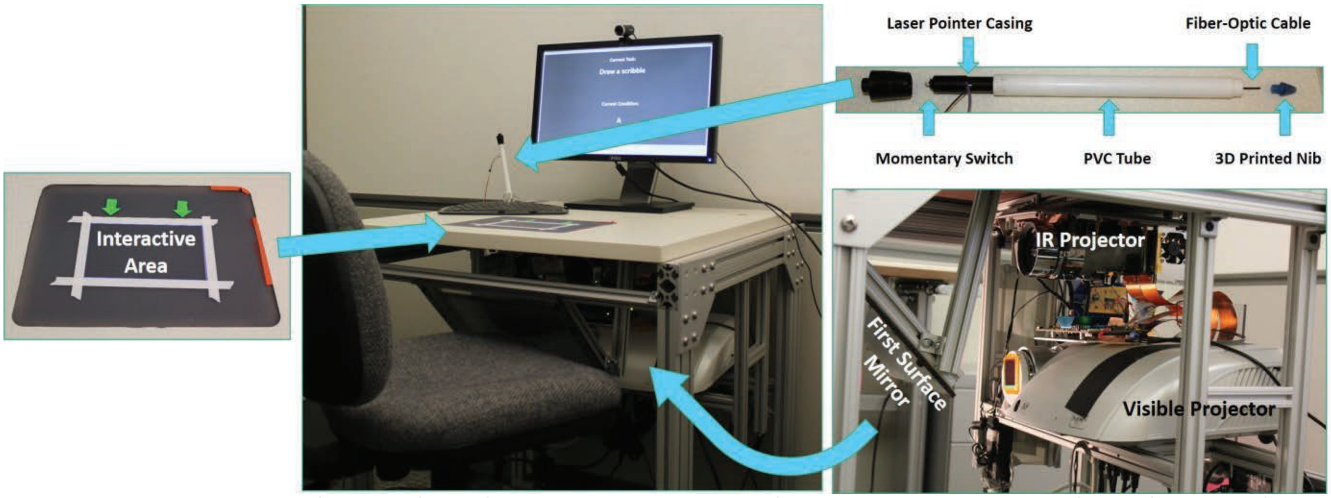

Figure 1: The prototype High Performance Stylus System, which is composed of two high-speed projectors, a first surface mirror for rear projection, and a fibre-optic based stylus. Further details about the hardware can be found in [24].

4.1 Participants

Twelve individuals (3 female) with normal or corrected-to-normal vision were recruited to participate in the study (M = 34, 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 experiment.

4.2 Equipment and Apparatus

To determine the minimum perceivable latency, the prototype High Performance Stylus System (HPSS) system from [24] was used (Figure 1). The HPSS employed two Texas Instrument Discovery 4100 high-speed projector kits [7], a first-surface mirror for rear-projection onto a diffuse surface, and a fibre-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 structured light technique [14]. The IR grey-coded patterns were projected at 17,000 frames per second at a resolution of 1920x1080. The patterns uniquely encoded every pixel in the image area. To provide visual feedback in the form of ink, the second Discovery 4100 kit refreshed at 23,000 binary frames per second, with a pixel resolution of 1920x1080. The stylus used a one-millimetre fibre optic cable to detect the grey-coded IR patterns. The fiber fit within a 3D printed 1.2-millimeter UV cured ABS plastic nib affixed to the end of a hollowed out laser pointer case. Opposite the nib was a momentary switch that 'activated' whenever the stylus was pressed against the screen. The stylus was close in weight and size to a Cintiq or Intuos stylus.

As operating system and application layers add latency to any system, the HPSS made use of the Discovery kits' on board Xilinx Virtex 5 FPGAs to decode the grey-code pattern sampled by the stylus and render feedback for the user via the visible projector. High-speed cameras and the method detailed in [23], in addition to timing on the FPGAs, verified the displayed latencies. As any camera-based approach introduces delays and noise, there is likely plus or minus half of a millisecond of error on the displayed latencies.

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 according to PEST. A custom C# and WPF program automatically determined and sent appropriate latency values to the system, gathered latency judgments from participants, and record the JND values for each task. A 21" Dell monitor provided participants feedback about the current task and condition and prompted them for their latency decision. 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.

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 our experiment. Participants were then informed that they would be performing a number of inking tasks and that we would be measuring the minimum latency that they could perceive during each task. 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 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.

4.4 Tasks

Three inking tasks were chosen based on their similarity to real world activities: 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

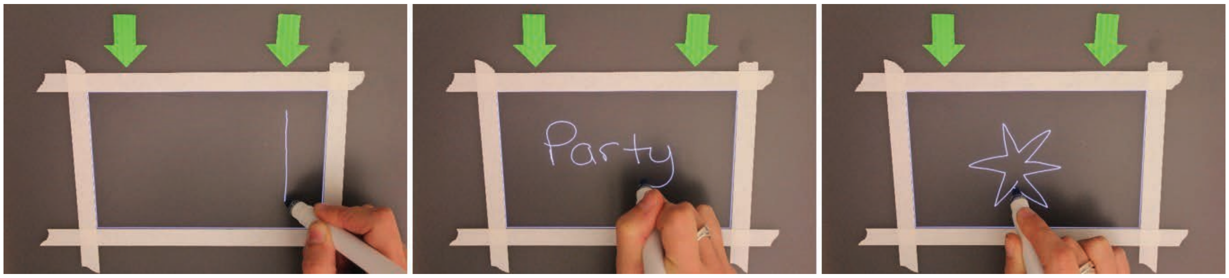

Figure 2: Tasks used in Experiment 2, from left to right: the line drawing, writing, and drawing tasks.

screen (Figure 2). 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, underline words, etc.). 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 2). ‘Party’ was used because it was required familiar, practised 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 from a single stroke or 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 the same character size across all trials.

In the *drawing* task, participants drew a six-sided star using one continuous stroke (Figure 2). 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 that 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. As we were interested in determining the absolute minimum latency users could perceive, we did not explicitly require participants 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 words, shapes, or stroke directions, participants would have had to complete another JND experiment. The experiment lasted approximately 30 minutes.

4.5 Results

As the JND thresholds were not normally distributed, a non-parametric Friedman’s ANOVA was performed with task as the only factor (i.e., line drawing, writing, and drawing; Figure 3). The ANOVA revealed that task did not have a significant effect on participant’s ability to perceive latency ($X^2(2) = 0.809$, $p = 0.667$, $\omega^2 = 0$). Participants were able to distinguish 7 and 53 milliseconds while line drawing (range: 31-76), 7 versus 50 milliseconds while writing the word ‘party’ (range: 32-87), and 7 versus 61 milliseconds while drawing the six-sided star (range: 21-82). 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.

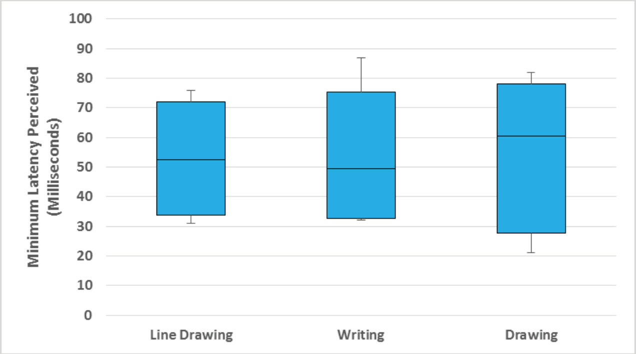

Figure 3: Median latency perceived while performing the writing, drawing, and line drawing tasks. The box boundaries represent the 5th and 95th percentiles.

4.6 Discussion

While performing everyday inking tasks, participants had a higher threshold for detecting latency than what has been found in prior work. For touch and simple box moving tasks, prior recommended target latencies were below 10 milliseconds [23] but approximately 40 milliseconds when scribbling an oscillating line [24]. 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 prior [24] 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, suggesting 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 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.

5 EXPERIMENT 2: ATTENTION AND VISUAL REFERENCE

As Experiment 1 suggested, a number of factors influence the perception of latency. We thus conducted another experiment to determine the extent that the visual and motor systems work together to aid in the perception of latency. Inspired by the differing judgement strategies reported in Experiment 1, we 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.

We were also interested in the effect of eliminating information about the motion of the stylus and hand, similar to indirect input scenarios. In such scenarios, interaction and visual attention is naturally decoupled, distributed along different planes or different 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. We thus manipulated the presence / absence of the hand within the visual field to mimic direct and indirect interaction scenarios.

5.1 Participants

Twelve naïve, right handed individuals (5 female) participated in the study (M = 33 years, range 24-44). Similar to the first experiment, all participants had normal or corrected-to-normal vision and had 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 participated in this study.

5.2 Equipment, Apparatus, and Procedure

The HPSS and procedure detailed in Experiment 1 was also used in Experiment 2.

5.3 Tasks

Four variations of the line drawing task from the first experiment were used (Figure 4). 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. 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 millimetres, 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 mimicked scenarios where the stylus is inaccurate. In the third condition, *large offset*, the ink was further offset to the left, approximately 65 millimetres. 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 enable a comparison with the last condition.

In the last condition, *hand not visible*, the ink was again offset 65 millimetres to the left of the nib but the hand was additionally obscured from view, using a foamboard flange placed vertically in the centre of the screen (Figure 4**).** 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.

Similar to the first experiment, 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.

5.4 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), $z = -2.86$, $p < 0.005$, $r = -0.58$ (Figure 5). The presence or absence of the hand and stylus is thus an important referent and influenced the perception of latency.

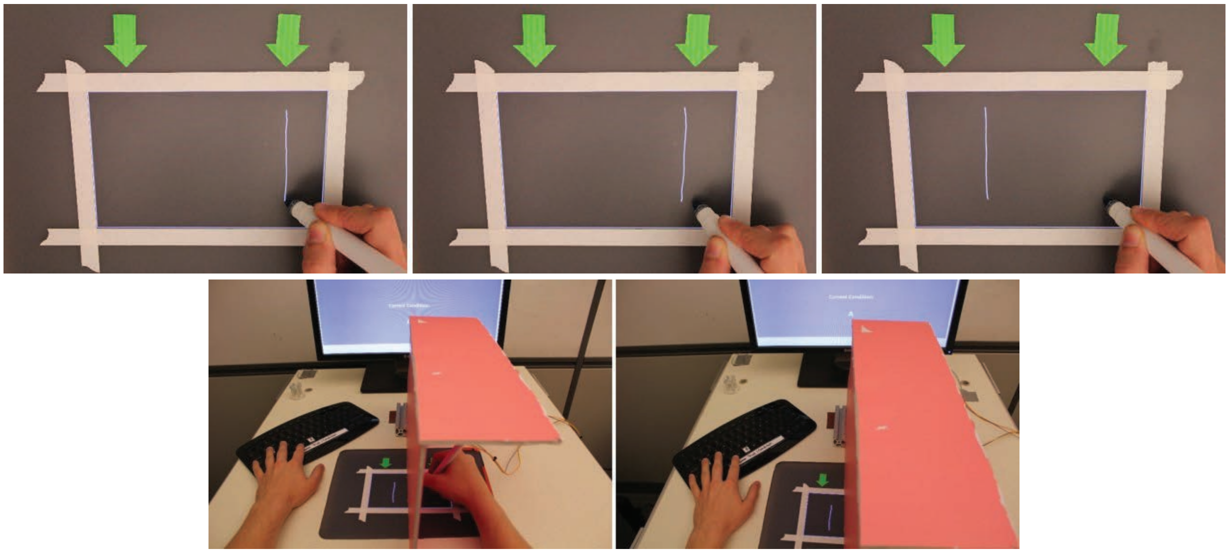

Figure 4: [Top] offsets used in Experiment 2, from left to right: no offset, small offset, and large offset. [Bottom] The hand not-visible condition. (Left) The participants placed their right arm and hand underneath the pink flange. (Right) The experiment from the perspective of the participant, wherein they could not see their hand or the stylus.

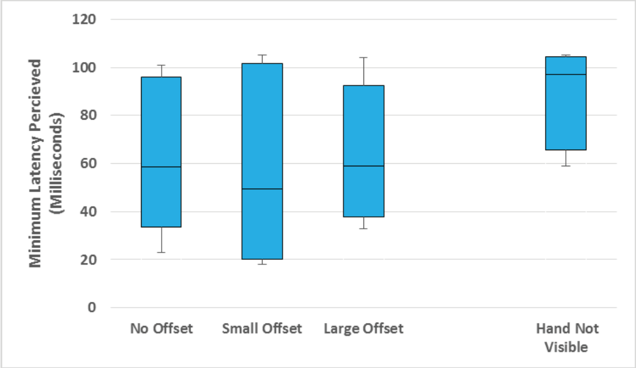

Figure 5: Median perceived latency while the ink was not offset, offset a small amount (i.e., 6.5 millimetres), offset to the opposite side of the screen (i.e., 65 millimetres), and while the hand was not visible. The box boundaries represent the 5th and 95th percentiles.

A Friedman's ANOVA examined the influence of the ink offset, (i.e., no offset, small offset, large offset). Offset distance did not have a significant effect on participant's ability to perceive latency ($X^2(2) = 4.167$, $p = 0.125$, $\omega^2 = 0.007$). Participants could distinguish 7 versus 59 milliseconds when no ink offset was present (range: 23-101), 7 versus 50 milliseconds when a small offset was present (range: 18-105) and 7 versus 59 milliseconds when a large offset was used (range: 33-104). Irrespective of the offset used, the median perceived latency was 55 milliseconds, similar to that found in Experiment 1. 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 judgements, instead relying on the relative movement of the stylus or hand.

5.5 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 our results and suggest the opposite. It is actually uncommon for participants to follow the nib continually with their eyes. Few participants actually 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 towards, as the stylus initiates interaction and the ink provides feedback about the action. Scanpath analyses have found that whenever the nib is located in the parafoveal or foveal region, it is often not attended to [2], 5, 8, 19, 29]. 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 [10, 28, 29, 30, 31]. 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 solely rely on the haptic feedback from their pen-wielding hand or the signals from their motor system to make latency judgements. 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 and increased latency discrimination. 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 no-hand 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 no-hand condition mimicked such a scenario quite well. These findings suggest that. 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, etc.).

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 to perceive latency, but this strategy is not supported by our 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 second experiment, we are not convinced of this. 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.

6 LATENCY PERCEPTION MODEL

While latency is simply the "delay between input action and the output response" [15], our previous and current explorations into latency have determined that the perception of latency is a complex, multi-faceted problem. In Experiment 2, 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 of other information streams, perhaps auditory or tactile cues from the stylus. These alternative data sources affected latency perception.

Based on our work and the prior literature, we have developed a model that describes the perceptual processes underlying latency perception in stylus and touch interaction. The model is composed of five elements: an input *action*, a *referent* stimulus, a *latency source*, output *responses*, and *contextual demands* (Figure 6). 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, etc.). This entity converts the input action into output responses and adds delay. The 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 our experiments demonstrated, changes in attention to and the location of responses can influence their use. Prior results using different sized boxes during dragging [13] 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 [13, 23]. 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 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).

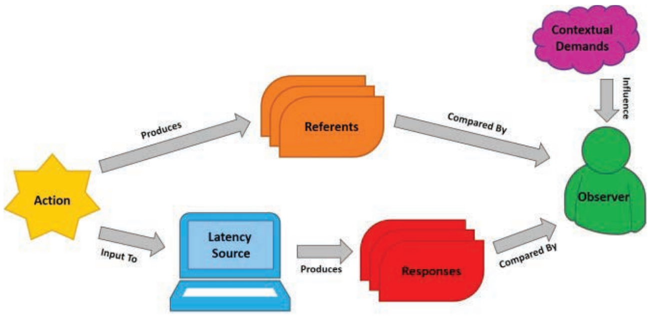

Figure 6: The Latency Perception Model, detailing the role of the input action, the resulting referent stimuli, the latency source, and the output responses. The observer compares the referents and responses when perceiving latency and is likely influenced by contextual demands such as task requirements, loci of attention, environmental factors, etc.

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, our experimentation determined that there are a variety of contextual demands influencing the observer. As suggested by Experiment 1, the judgement 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 we cannot make definitive comparisons, tasks that require more attention (inking versus moving a box) and increased cognitive load seem to redirect resources away from perceiving the referents and responses, making latency judgements more difficult.

Although little is known about latency perception, the model 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.

7 FUTURE WORK

A great deal of attention has been devoted towards latency perception recently, but there still is a great deal open for exploration. Miller's 100 millisecond latency hypothesis focused on the issue of latency tolerance whereas our experimentation examined perceptual thresholds [20]. 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. Our experiments determined the display latency that can be perceived but not the effects of delayed or slow sampling. While we advocate for decreasing latency in the whole pipeline, it remains to be seen how perception and user satisfaction would change if devices rendered quickly but stroke completeness and accuracy were unpredictable.

As reducing the actual system latency requires many incremental improvements throughout the data pipeline, it may be worthwhile to use additional information to modulate the processing devoted to latency reduction whenever the user is looking at the display or the stylus is about to touch the screen. The use of eye tracking and 'pre-touch' information from the stylus or finger is also an interesting avenue of research. By predicating and anticipating strokes or actions before contact is made with the screen, it may be possible to decrease the perceived latency inherent in a system.

8 CONCLUSION

In our experiments, we dove further into understanding the basics of latency perception by performing simple manipulations on the tasks performed and the presence and locations of visual reference cues. By offsetting the location between the ink and stylus and removing the hand and stylus from the visual field in one condition, we determined that low-latency judgments are largely visual, using the relative movement between the stylus and ink for an accurate judgment. Participants were unable to distinguish latencies below 97 milliseconds when the hand and stylus were not visible. We additionally found that latency perception is task dependant, with inking tasks degrading one's ability to discern between various levels of latency when compared to tasks used in prior work. These results informed the Latency Perception model, a model focusing 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.

REFERENCES

[1] B.D. Adelstein, D.R. Begault, M.R. Anderson, and E.M. Wenzel. Sensitivity to Haptic-Audio Asynchrony. In *Proc. of Int. Conf. on Multimodal Interfaces*, 73-76, 2003.

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

[3] M. Annett, F. Anderson, W.F. Bischof, and A. Gupta. The Pen is Mightier: Understanding Stylus Behaviour While Inking on Tablets. In *Proc. of GI*, 2014.

[4] G. Anderson, R. Doherty, and S. Ganapathy. User Perception of Touch Screen Latency. *Design, User Experience, and Usability Part 1*, 195-202, 2011.

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

[6] S. Dahl and R. Bresin. Is the Player More Influenced by the Auditory than the Tactile Feedback from the Instrument? In *Proc. of the COST-06 Workshop on Digital Audio Effects,* 194-197, 2001.

[7] DLP Discovery 4100 Development Kit. http://www.ti.com/lit/an/dlpa008/dlpa008.pdf. Accessed July 2013.

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

[9] A. Freed, A. Chaudhary, and B. Davila. Operating System Latency Measurement and Analysis for Sound Synthesis and Processing Applications. In *Proc. of Int. Computer Music Conf.*, 479-481, 1997.

[10] E. Gowen and R.C. Miall. Eye-tracking Interactions in Tracing and Drawing Tasks. *Human Movement Science*, 25: 568-585, 2006.

[11] A. Henzen, N. Ailenei, F. Di Fiore, F. Van Reeth, and J. Patterson. Sketching with a Low-Latency Electronic Ink Drawing Tablet. In *Proc. of 3rd Int. Conf. on Computer Graphics and Interactive Techniques in Australasia and South East Asia*, 51-60, 2005.

[12] A. Henzen, N. Ailenei, F. Van Reeth, G. Vansichem, R.W. Zehner, and K. Amundson. An Electronic Ink Low Latency Drawing Tablet. *SID Symposium Digest of Technical Papers*, 35(1): 1070-1073, 2004.

[13] 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.

[14] J.C. Lee, P. Dietz, D. Maynes-Aminzade, R. Raskar, and S. Hudson. Automatic Projector Calibration with Embedded Light Sensors. In *Proc. of UIST*, 123-126, 2004.

[15] I.S. MacKenzie and C. Ware. Lag as a Determinant of Human Performance in Interactive Systems. In *Proc. of CHI*, 488-493, 1993.

[16] T. Maki-Patola. Musical Effects of Latency. *Suomen Musiikintutkijoiden*, 9: 82-85, 2005.

[17] T. Maki-Patola and P. Hamalainen. Latency Tolerance for Gesture Controlled Continuous Sound Instrument without Tactile Feedback. In *Proc. of Int. Computer Music Conf.*, 1-5, 2004.

[18] N.A. McMillian and D.C. Creelman, *Detection Theory: a User's Guide*, Cambridge University Press, Cambridge, 1991.

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

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

[21] M. Montag, S. Sullivan, S. Dickey, and C. Leider. A Low-Cost, Low-Latency Multi-Touch Table with Haptic Feedback for Musical Applications. In *Proc. of Int. Conf. on New Interfaces for Musical Expression*, 8-13, 2011.

[22] A. Ng and P. Dietz. The Need for Speed in Touch Systems. *SID Symposium Digest of Technical Papers,* 2013.

[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] A. Ng, M. Annett, P. Dietz, A. Gupta, and W.F. Bischof. In the Blink of an Eye: Investigating Latency Perception during Stylus Interaction. In *Proceedings of CHI 2014*.

[25] A.A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis. From Remote Media Immersion to Distributed Immersive performance. In *Proc. of ACM SIGMM Workshop on Experiential Telepresence*, 110-120, 2003.

[26] S. Seow. *Designing and Engineering Time: The Psychology of Time Perception in Software*, Pearson Education, Boston, 2008.

[27] M.M. Taylor, and C.D. Creelman. PEST: Efficient Estimates on Probability Functions. *Journal of the Acoustical Society of America*, 41: 782-787, 1967.

[28] J. Tchalenko. Eye Movements in Drawing Simple Lines. *Perception*, 36: 1152-1167, 2007.

[29] J. Tchalenko and R.C. Miall. Eye-hand Strategies in Copying Complex Lines. *Cortex*, 45:368-376, 2009.

[30] N. Toyoda, R. Yamamoto, and T. Yubuta. Clarification of Transition Conditions for Eye Movement While Generating a Trajectory Using the Upper Limb. *SI International*, 106-111, 2011.

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

[32] D. Vogel and R. Balakrishnan. Direct Pen Interaction with a Conventional Graphic User Interface. *Human Computer Interaction*, 25: 324-388, 2010.

[33] A. Wald. *Sequential Analysis*, John Wiley & Sons Inc. New York, 1947.

[34] J. Wright and E. Brandt. System-Level MIDI Performance Testing. In *Proc. of Int. Computer Music Conf.*, 318-331, 2001.