Steadied-Bubbles: Combining Techniques to Address Pen-Based Pointing Errors for Younger and Older Adults

Karyn Moffatt and Joanna McGrenere
Department of Computer Science, University of British Columbia
201-2366 Main Mall, Vancouver, BC V6T 1Z4
{kmoffatt, joanna}@cs.ubc.ca

ABSTRACT
Tablet PCs are gaining popularity but many older adults still struggle with pointing, particularly with two error types: *missing*, landing and lifting outside the target bounds; and *slipping*, landing on the target, but slipping off before lifting. To solve these problems, we examined the feasibility of extending and combining existing techniques designed for younger users and the mouse, focusing our investigation on the Bubble cursor and Steady Clicks techniques. Through a laboratory experiment with younger and older adults, we showed that both techniques can be adapted for use in a pen interface, and that combining the two techniques provides greater support than either technique on its own. Though our results were especially pertinent to the older group, both ages benefited from the designs. We also found that technique performance depended on task context. From these findings we established guidelines for technique selection.

Author Keywords
Pen-based interaction, pointing facilitation, older adults, error prevention.

ACM Classification Keywords
H5.2. Information interfaces and presentation: User interfaces—*input devices and strategies*.

General Terms
Design, Experimentation, Human Factors.

INTRODUCTION

Technology is increasingly being promoted as a means of addressing age-related cognitive and sensory impairments and enabling seniors to live more independently (e.g., [14, 15, 19, 21]). Pen-based devices such as Personal Digital Assistants and Tablet PCs are appealing platforms for these endeavors because they are small, mobile, and powerful. They also allow users to take full advantage of their hand-eye coordination skills in a familiar form of interaction [8]. When compared to a mouse, pen input has been shown to be particularly beneficial for older adults [5].

CHI 2010, April 10–15, 2009, Boston, Massachusetts, USA.

Unfortunately, many older adults encounter difficulties when using a pen for input [18]. As with all input devices, selection errors can be costly and overly frustrating [4]. Although most modern programs offer extensive undo functionalities, these do not necessarily address all costs associated with making an error. For example, selecting the wrong item from the Windows Start menu is easily corrected by closing the undesired program and reselecting the intended one, but the user must first wait for the unwanted program to load, which can be time-consuming. Techniques that reduce errors can have a large positive impact for those users who, like many older adults, are easily confused and discouraged by them. This is often lost in laboratory studies as it is impossible to encapsulate the true cost of recovery time even when penalties are included.

In a previous study [18], we identified two main sources of general pen-based target acquisition difficulty for older adults: (1) *Missing*, landing and lifting outside the target bounds; and (2) *Slipping*, landing inside the target bounds, but slipping out before lifting the pen. (In point and tap interaction, selection is determined by the location of the pen at lift off.) Missing was common to both older and younger adults and remained relatively constant across age. Slipping, in contrast, was unique to the older users and accounted for almost half of the errors for that group.

To address the needs of older individuals using a pen, we explored the feasibility of extending and combing existing techniques for younger users and mouse interaction. Specifically, we focused on two promising mouse-based techniques: Steady Clicks [23] and Bubble cursor [10]. Furthermore, as these two techniques address different aspects of target acquisition, we also assessed the feasibility of combining them to capitalize on each of their strengths.

Steady Clicks

Steady Clicks is a mouse-based technique designed to help in situations where the user successfully clicks down on a target but slips off before releasing [23]. It works by freezing the cursor at the button down location until either it is released (causing a steadied click to occur), or the mouse is moved beyond the freeze threshold (returning it to normal operation). An evaluation of Steady Clicks found that it enabled motor-impaired participants to select targets using significantly fewer attempts, and for those with the highest slip rates, to select them significantly faster [23].

Steady Clicks was originally designed to address mouse-based slipping errors. Although slipping is common to both mouse and pen interaction, with a mouse, it is generally attributed to an inability to hold the mouse still while clicking [23]. Tap selection does not have an analogous button clicking action, so it is not immediately clear that techniques designed to reduce slipping for the mouse will directly translate to pen interfaces. One potential barrier to using Steady Clicks with a pen is that it alters the ratio between mouse and cursor movement. The direct mapping between the cursor and pen tip makes this less ideal.

Bubble Cursor

The Bubble cursor is a dynamic area cursor [26], in which a circular cursor grows and shrinks to capture the nearest (and only the nearest) target [10]. Prior evaluation showed that it was faster and more accurate than a standard point cursor, and that its performance could be predicted by Fitts' Law by using the size of the cursor as the target's effective width [10]. Although this technique was not designed to address slipping, it essentially makes targets bigger in motor space, which should reduce the likelihood of a slip movement resulting in an error. Bubble cursors have not been evaluated with older adults, but the static area cursors upon which they are based have been shown to improve mouse-based pointing performance for older adults [26], suggesting promise.

Combining Approaches: The Steadied-Bubble

Bubble cursor and Steady Clicks each target different aspects of pointing. The Bubble cursor mostly helps ease the initial positioning of the cursor, while Steady Clicks is designed to help keep it steady once it is in place. Thus, it seems feasible to combine them into a single technique that fully captures the advantages of each. For our combined Steadied-Bubble approach, a circular Bubble cursor grows and shrinks to capture the nearest target while the pen-tip is within the hover-range of the display. Once the pen-lands on the screen, the cursor is frozen in both its location and size; that is, it is locked onto the last target captured before landing. If the pen moves beyond the freeze threshold, the Bubble cursor returns to its normal operation: the center of the area cursor tracks the tip of the pen, and the cursor grows and shrinks to capture the nearest target. Note that the freeze threshold is constant, but for any particular freeze, it may be larger or smaller than the radius of the Bubble cursor, depending on the target layout and density.

Each of the Steady Clicks, Bubble cursor and Steadied-Bubble approaches has inherent benefits and drawbacks. The Steady Clicks approach is cognitively simpler, but the mismatch between the cursor's position during freezing and the physical pen tip may be confusing to some users. Another disadvantage of this approach is that if the user misses the target on landing, it is harder to correct the selection by sliding the pen along the surface, as the user must first break the freeze threshold. An advantage of the Bubble cursor is that it could potentially address both slipping and missing. However, its effectiveness degrades as target density increases, making it least helpful when it is most needed. That is, when targets are dense, errors are more likely to activate unwanted functionality, and such errors are more costly than selections on inactive space. Another disadvantage of the Bubble Cursor technique is that it depends on knowing the location of all targets, which is not always possible. The combined Steadied-Bubble cursor seems most promising. The strong visual feedback provided by the Bubble cursor should help ease the mismatch caused by freezing, and overall it should offer the most support. However, it is more complicated than either technique on its own, which some users might find overwhelming.

To evaluate these tradeoffs, we conducted a laboratory experiment comparing the Bubble, Steady, and Steadied-Bubble cursors to each other and to standard point and tap, with 12 younger and 12 older adults. We found empirical evidence demonstrating that (1) Bubble was effective at reducing both slips and misses, but only when targets were not directly adjacent; (2) Steady was only effective at reducing slips, but its support was independent of target spacing; and (3) combining them into a single technique, Steadied-Bubble, successfully integrated the benefits of each—Steadied-Bubble prevented misses when targets were not adjacent, and slips independent of spacing.

The main contribution of this work is these empirical results. They demonstrate that both the Bubble cursor and Steady Clicks techniques can be successfully adapted for use in a pen-based interface, and that they are particularly helpful for older adults. They further establish that these techniques can be successfully integrated to address multiple sources of target acquisition difficulty across a range of task contexts. Finally, we draw on these context-specific findings to establish guidelines for techniques selection.

RELATED WORK

We begin our coverage of the literature with an overview of the general effects of aging on motor skill to highlight the reasons for age-related differences in targeting ability. We then describe research investigating novel pen techniques.

There is a considerable body of literature that has examined the negative effects of aging on the aspects of motor control that pertain to general targeting ability, both with respect to mouse use and interaction in the physical world. Research has found that older adults use different strategies for the speed-accuracy tradeoff involved in movement control. Older adults tend to be more conservative, and make more corrective submovements once inside the target [24]. They have also been found to cover less distance with their primary movement [13], to make more submovements en route [12], to make less smooth movements [27], and to have difficulty staying on the target while clicking [22]. In addition, slower selection speeds have been attributed to lower peak velocities [12, 13], longer deceleration phases [13], and more pauses while homing in on the target [12].

The most extensive evaluation of pen-based interaction to date is work by Ren and Moriya [20]. They constructed a state transition model for pen-based target selection and using that model derived six selection strategies. They found that for targets smaller than 1.8 mm, Slide Touch (selection at the moment the pen-tip first touches a target after landing) was best in terms of speed, accuracy, and preference. However, they cautioned that this technique would not be suitable for dense displays, for which they recommended either Direct On (which relies on the pen landing on the target), or Direct Off (standard point and tap). Both require good hand-eye coordination; thus, it is unclear whether they are suitable for older users.

Beyond pointing, other researchers have investigated alternate mechanisms that may be better suited to pen input. Mizobuchi and Yasumura compared tapping to circling for a multi-target selection task [17]. They hypothesized that circling would be faster and more accurate than tapping, but found that it was only better in the specific situation where multiple targets formed a cohesive group with low shape complexity. Accot and Zhai compared tapping to crossing and found crossing was at least as fast and had similar accuracy [1]. Though not outright better, they suggested there may be specific situations in which crossing has advantages. Others have expanded on their ideas; for example, by investigating fluid multi-action sequences [2], and their use for supporting motor-impaired users [25].

To date, very little work has examined the use of pen-based interaction with older adults. Charness et al. performed an age-related comparison of the mouse and the light-pen [5]. They found that the pen outperformed the mouse for all ages and reduced the performance gap between ages, but that the mouse was rated as being more acceptable and easier to use (across ages). However, this work was done with a light-pen on a vertical monitor, which required the pen to be held up unnaturally. Modern Tablet PCs are designed to be more comfortable, and thus, should result in higher satisfaction.

More recently, Hourcade and Berkel compared two pen-based selection techniques, tapping and touching (selection if the pen touches the target at any time before tap up), across three adult age groups, and found that for the smallest target size examined (3.8 mm) the oldest group was more accurate using touch, but found it more tiring [11]. One limitation of this technique is that when targets are directly adjacent, it degrades to tap.

EXPERIMENTAL METHODOLOGY

We conducted a controlled laboratory experiment with younger and older adults to compare the individual and combined effectiveness of the Bubble cursor and Steady Clicks techniques for reducing pen-based targeting errors.

Apparatus

For the experiment, we used a Wacom Cintiq 12WX pen tablet, and a 2.26GHz Duo Core laptop with 2 GB of RAM and Microsoft Windows XP. A Cintiq Classic pen was used for all the tasks, with the barrel buttons deactivated to ensure participants did not use them unintentionally. The Cintiq senses 1024 levels of pressure and has a 307 mm (12.1 inch) diagonal display with a resolution of 1280×800 pixels (261×163 mm; i.e., pixel width = 0.2 mm.) The software was coded in Python using the Pygame SDK and the Wintab wrapper of the Python Computer Graphics Kit; it recorded all timing and error data. The Cintiq was inclined to 25 degrees from horizontal using its built-in stand, to position it at a comfortable viewing angle. Participants were encouraged to adjust the position of their chair and the computer for comfort and most did.

Pointing Techniques

We examined the following four cursor types *(CT)*.

- **Control.** Standard arrow cursor. No slip filtering.
- **Bubble.** Dynamic bubble cursor. No slip filtering.
- **Steady**. Standard arrow cursor. Slips filtered.
- **Steadied-Bubble.** Dynamic bubble cursor. Slips filtered.

Steady and Steadied-Bubble filtered movement below a threshold of 60 px (12 mm). That is, while the pen remained within 60 px of its initial landing position, the cursor remained fixed at this position, and lifting the pen resulted in a selection event at the initial landing position. Once the pen moved more than 60 px away, the cursor returned to normal operation, tracking the tip of the pen; lifting resulted in selection at the lift position. The 60 px threshold was chosen to be larger than most of the slips observed in [18]. For Bubble and Steadied-Bubble, the cursor was rendered in a light semitransparent grey. We enforced a maximum diameter of 100 px (20 mm) based on the suggestion by Grossman and Balakrishnan [10].

Participants

We recruited 12 participants from each *age group (AGE)*, for a total of 24 participants.

- **Younger.** Aged 19–29 (M = 23, 5 female, 7 male)
- **Older.** Aged 65–86 (M = 73, 6 female, 6 male)

The younger participants were recruited through campus postings. They received $15 for participating and completed the study in 60–80 minutes (M = 68 min). The older participants were recruited though community postings and word-of-mouth advertisement. On average, they took longer than the younger participants (completing the study in 75–120 minutes, M = 88 min), and received $20 for their participation. Participants were right-handed and free of diagnosed motor impairments to their hands. Additionally, they all had normal or corrected-to-normal eyesight.

To control for biases between age and Tablet PC experience, all were novices to pen-based computing. Within and across each age group, participants had a wide range of computer experience. Nonetheless, there were some notable differences: younger participants were more frequent users, used a greater number of applications, and were familiar with a greater number of advanced tasks. Their self-rating of expertise was also higher. However, the older participants had been using computers for longer.

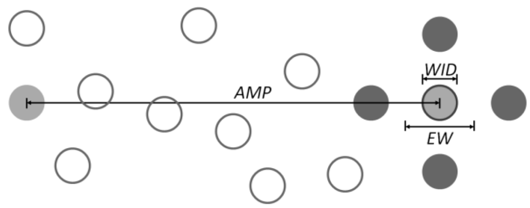

Figure 1. Experimental setup. In this figure, light grey denotes the previous and current goal targets, with the current goal indicated by a dark outline. The rest are distracters, with the *neighbors* filled in dark grey. (*DEN* = 0.5, *EWR* = $^{EW}/_{WID}$ = 2.)

Task

Our task is shown in Figure 1, and was modeled after that used by Grossman and Balakrishnan [10]. Participants selected a series of goal targets, appearing at unpredictable locations. To control the inactive whitespace around the goal target, four distracter targets were equidistantly placed around the goal, in the line of and perpendicular to the axis of approach. Additional distracter targets were placed in the scene to create varying levels of overall target density.

The goal target was rendered as a solid green circle, and distracter targets as grey outlined circles of the same size. Visual feedback was provided by changing the appearance of targets: (1) when the pen tip hovered over a goal target it turned a deep red (distracter targets turned a solid dark grey), and (2) when the pen tip touched a goal target it turned a brighter red with a dark red border (distracter targets similarly turned a lighter grey with a dark grey border). We provided the latter form of feedback to help participants determine how much pressure was needed.

Consistent with Grossman and Balakrishnan's study design [10], we varied the following factors.

Target Width (WID) specifies the diameter of the goal target. We used three target widths for the experiment: 12, 24, and 36 pixels (2.4, 4.8, and 7.2 mm). These sizes are in line with previous studies (e.g., [10, 11, 20]), and roughly correspond to the following common widgets: the height of a text link, the size of a small toolbar icon (or the height of a menu item), and the size of a larger icon (e.g., the height of Windows Start menu items using the 'large icon' option).

Amplitude (AMP) is the distance to the goal target from the starting position of the trial (i.e., from the previous trial's goal). We examined three amplitudes: 256, 382, and 512 pixels (51, 76, and 102 mm) to explore a range of distances.

Effective Width Ratio (EWR) specifies the amount of inactive whitespace surrounding the goal target (i.e., the distance to its *neighbors*). Target spacing is particularly important for Bubble and Steadied-Bubble. Thus, we express it as the ratio between the *effective target width* (for Bubble and Steadied-Bubble) and the actual target width (i.e., *EWR* = $^{EW}/_{WID}$). We used three values for this factor: 1, 2, and 3. When *EWR*=1, the goal target is directly adjacent with its neighbors.

Distracter Density (DEN) refers to the number of other targets on the screen. We used the same three levels as in [10]: 0, 0.5, and 1. *DEN*=0 reflects no distracter targets (except for the four neighbors), *DEN*=0.5, a moderate target density (see Figure 1), and *DEN*=1, a high target density. For complete details on the distracter placement, see [10].

At the start of each new cursor condition, participants were introduced to the cursor and given 10 practice trials. Participants then completed four blocks of trials with each cursor, with a short break between blocks. Each block consisted of 81 trials representing one of each possible combination of *WID*, *AMP*, *EWR*, and *DEN*, for a total of 324 trials per cursor condition. The order of presentation of trials was consistent with [10] and was as follows. Each combination of *WID*, *EWR*, and *DEN* was presented in a random order. Within each of these combinations, all three levels of *AMP* were presented together (in a random order). This was done to provide some sense of continuity between trials. Our early pilot runs of the experiment also found that having all four factors change every trial was disorientating.

Design

We used a 2×4×3×3×3×3 mixed factorial design, with *AGE* as a between-subjects factor, and *CT*, *WID*, *AMP*, *EWR*, and *DEN* as within-subjects factors. Cursor type (*CT)* was a within-subjects factor to increase the power of the design. Each participant was assigned to one of four presentation orders, following a balanced Latin square.

Procedure

The experiment was designed to fit into a single 120-minute session. We began with a series of standardized tests of sensory-perceptual and motor skills. Next, participants were asked to complete a brief questionnaire about their background and computer experience. They were then introduced to using the pen-based device, and shown that (1) they can rest their hand on the screen during input, (2) the computer tracks the pen both when it is touching the screen and when it is slightly above it, and (3) an onscreen cursor provides feedback of the current cursor location. Once participants were comfortable using the pen, the Cintiq was calibrated using the built-in utility.

Participants then completed the experimental tasks. Following each condition, participants completed a short questionnaire about that condition. Between conditions, participants completed short verbal distracter tasks. These tasks were chosen to engage participants mentally, but not physically to rest their arms. At the end of the experiment, participants were asked to rank the interfaces on a number of factors and encouraged to make additional comments.

Measures

For accuracy, we measured errors individually as the total number of slips and the total number of misses in each condition. We additionally included trial time as a measure, to provide an overall indication of performance. For time, the median was used to reduce the influence of outlier trials, and an implicit error penalty (participants could not advance to the next trial until they correctly completed the current trial) was used to discourage participants from overly focusing on speed. However, we note that while it is helpful in terms of motivating a balance between speed and accuracy, this penalty underestimates the true cost of errors as it treats errors on distracter targets the same as errors on inactive whitespace. In real-world tasks, selection of an unwanted feature typically requires additional corrective action. After each condition participants gave Likert scale ratings and at the end of the study, they ranked the cursor techniques on speed, ease, frustration, and preference.

Motivation

To motivate quick and accurate performance, an additional \$10 incentive was awarded to the top 1/3 performers in each age group. The 1/3 ratio was chosen to encourage participants to believe they had a reasonable chance of succeeding. To help participants gauge their performance, graphical feedback was presented during the breaks between blocks. This feedback included a graph of the participant's speed for all blocks completed with that cursor, and a text summary of total time and errors for the most recent block.

Hypotheses

Our hypotheses are all relative to Control.

H1. *Bubble will reduce both slips and misses, but only when the surrounding targets are not directly adjacent.* That is, when *EWR*=1, there will be no difference between Bubble and Control, but when *EWR*>1, Bubble will result in both fewer slips and fewer misses.

H2. *Independent of target spacing, Steady will reduce slips, but it will not affect misses*. That is, we predict that for all *EWR*, Steady will result in fewer slips, but not fewer misses, than Control. Combined with H1, Bubble will reduce more errors overall than Steady, for *EWR*>1.

H3. *Steadied-Bubble will reduce slips when targets are directly adjacent, and both slips and misses when they are not.* That is, it will fully integrate the individual benefits of Bubble and Steady. Correspondingly, it will be the most effective technique at reducing errors.

H4. *The experimental cursors (Bubble, Steady, and Steadied-Bubble) will reduce total errors for both age groups, but the impact will be larger for the older group as they will benefit from both slip and miss reduction*. This hypothesis is based on our previous finding that older users make both slip and miss errors, whereas younger users predominantly miss [18].

H5. *A greater proportion of errors in Bubble and Steadied-Bubble will land on a distracter target*. Although Bubble and Steadied-Bubble will both result in fewer errors than Control, when errors do occur, they will be more likely to hit a distracter target. For Steady, we do not predict an increase in distracter hits.

RESULTS

For each of our main measures (trial time, misses and slips), we performed a repeated measures ANOVA. For trial time, we performed a full analysis across all factors. For misses and slips, we focused our analysis on just those factors for which we had hypotheses (*AGE*, *CT*, *WID*, *EWR*), collapsing across the other factors. Initial analysis of the data did not suggest any main or interaction effects for *AMP* and *DEN*, and with only four trials (per each combination of all factors), the error data was too sparse to measure the differences in which we were interested.

In our reporting of F-statistics, where *df* is not an integer, we have applied a Greenhouse-Geisser adjustment for non-spherical data. All pairwise comparisons were protected against Type I error using a Bonferroni adjustment. Along with statistical significance, we report partial eta-squared (η^2), a measure of effect size. Roughly speaking, .01 is considered a small effect, .06 medium, and .14 large [7].

Errors

Table 1 provides a summary of the error rates for the younger and older groups. Consistent with other research [11, 20], the majority of errors in this study occurred on the smallest target (*WID*=12). For the larger widths, error rates were low and skewed towards zero, suggesting a floor effect. Skewed data can invalidate the results of an ANOVA analysis; thus, we focus our statistical analysis on just the trials with *WID*=12. We do note, however, that while the largest differences, and correspondingly the most practically significant ones, occur at *WID*=12, a similar but highly attenuated pattern is evident for the other widths, as shown in Figure 2.

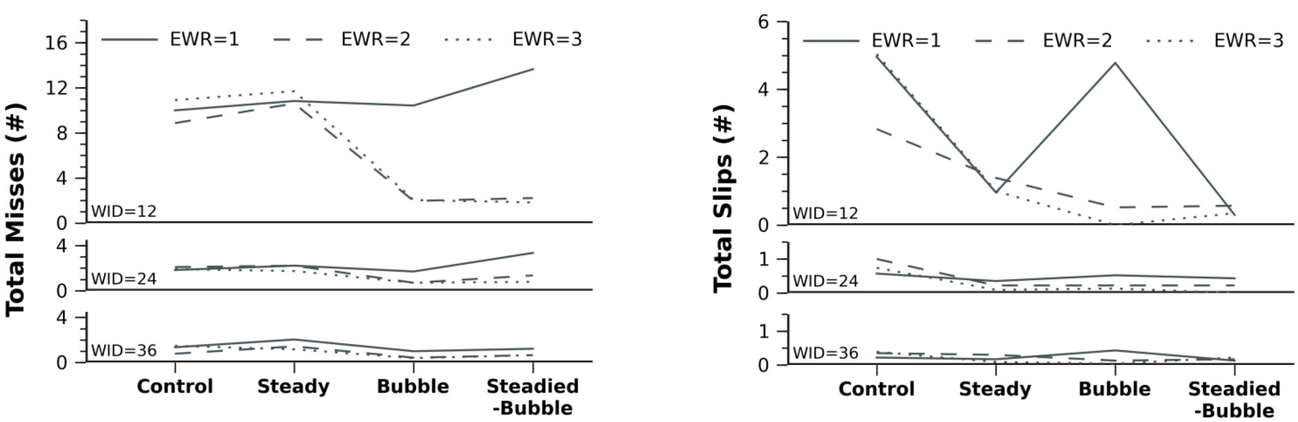

Figure 2. Average total misses (left) and slips (right) by *CT, WID,* and *EWR* (*N*=23, for 36 trials).

		Younger		Older	
		M(*SD*)	Range	*M*(*SD*)	Range
Control	Misses	7.5(6.1)	3.1–24.4	16.3(9.5)	4.3–31.2
	Slips	3.9(3.0)	0.3–11.1	5.9(3.4)	0.6–11.1
	Total	11.5(7.4)	4.0–28.7	22.2(11.7)	7.4–41.0
Steady	Misses	8.5(8.8)	2.2–33.6	18.2(12.3)	6.2–45.7
	Slips	1.7(2.5)	0.3–9.0	1.1(0.6)	0.3–1.9
	Total	10.2(9.6)	2.5–35.2	19.3(12.6)	6.8–47.5
Bubble	Misses	3.4(2.3)	0.6–8.0	8.3(4.4)	2.8–16.4
	Slips	1.9(2.0)	0.0–7.7	2.3(1.5)	0.6–6.2
	Total	5.2(3.3)	1.5–12.0	10.6(4.9)	4.0–19.4
Steadied -Bubble	Misses	5.1(3.7)	1.5–14.5	10.5(5.6)	3.4–19.8
	Slips	1.1(2.4)	0.0–8.3	0.4(0.9)	0.0–2.5
	Total	6.2(5.0)	1.5–16.0	10.9(6.0)	3.4–19.8

Table 1. Overall error rates by *AGE* and *CT* (*N*=24). These numbers are based on all 324 trials per cursor condition.

Even after filtering out the larger widths, some of the error rates were floored. This is not entirely unfortunate; all cases corresponded to instances where an experimental cursor substantially reduced one of the error types (i.e., none of the results for Control are floored). However, to ensure that these measures do not bias the statistical results, we additionally relied on confidence intervals to aid our interpretation of the ANOVA results. Specifically, we only report those significant pairwise comparisons between cursors from the ANOVA, where confidence interval analysis also found a significant difference.[1]

One participant in the younger group had unusually high error rates. His performance was outside the 1.5 interquartile range and more than two standard deviations from the mean. Although analysis with and without him yields the same conclusions, we exclude him to better reflect the performance of the younger group as a whole.

Miss Errors, WID=12

Bubble and Steadied-Bubble significantly reduced misses when targets were not adjacent. There was a significant main effect of *CT,* and as shown in Figure 3, a significant interaction between *CT* and *EWR* (*CT*: $F_{1.9,40.1} = 8.32$, $p = .001$, $\eta^2 = .284$; *CT* × *EWR*: $F_{3.9,82.0} = 11.0$, $p < .001$, $\eta^2 = .344$). Pairwise comparisons revealed that Bubble and Steadied-Bubble both resulted in significantly fewer misses than Control and Steady for *EWR*=2 and *EWR*=3 (all $p < .005$), but not for *EWR*=1. Steady was not significantly different from Control for any *EWR*.

The older adults missed significantly more. There was a main effect of *AGE* ($F_{1,21} = 15.5$, $p = .001$, $\eta^2 = .425$). On average, the older group missed 2.67 times for every miss by the younger group (older: M = 135.7, SD = 65.2; younger: M = 50.7, SD = 30.6; per all 1296 trials).

[1] A confidence interval is an indication of the reliability of a measured estimate. It is more conservative than an ANOVA analysis because it does not pool variances; thus, a floor effect in one level of measurement does not affect the confidence intervals of other levels of measurement. To aid the reader, 95% confidence intervals are included as error bars in all our graphical results. Nonoverlapping error bars represent significantly different results.

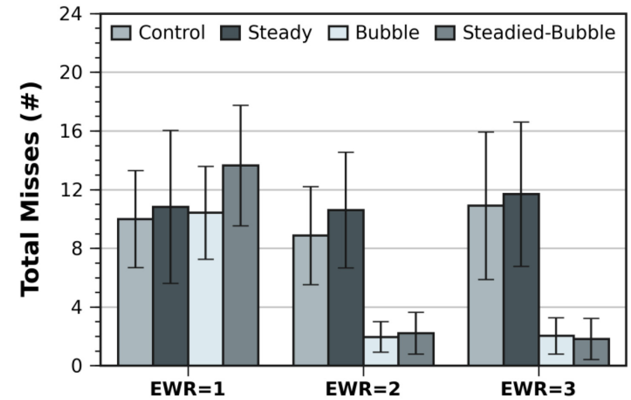

Figure 3. Average total misses (for 36 trials) by *CT* and *EWR* (*WID*=12, *N*=23). Note the significant drop in misses for Bubble and Steadied-Bubble for *EWR*=2, 3.

There was also a main effect of *EWR* and a significant interaction between *AGE* and *EWR* (*EWR*: $F_{2,42} = 39.8$, $p < .001$, $\eta^2 = .655$; *AGE* × *EWR*: $F_{2,42} = 5.24$, $p = .009$, $\eta^2 = .200$). These results simply mirror the other findings. For *EWR*=1, both groups incurred roughly twice as many misses as they did for *EWR*=2 and *EWR*=3 because Bubble and Steadied-Bubble were not effective at reducing errors at *EWR*=1. Moreover, because the older adults missed more in general, this doubling at *EWR*=1 resulted in a greater increase for them, which explains the interaction.

Slip Errors, WID=12

Steady and Steadied-Bubble reduced slips and performed consistently across target spacings. Bubble also reduced slips, but only when targets were not adjacent. There was a significant main effect of *CT* and a significant interaction between *CT* and *EWR* (*CT*: $F_{1.4,29.8} = 29.2$, $p < .001$, $\eta^2 = .582$; *CT* × *EWR*: $F_{2.9,62.6} = 15.5$, $p < .001$, $\eta^2 = .425$).

Pairwise comparisons of the *CT* × *EWR* interaction (shown in Figure 4) revealed that Bubble resulted in significantly fewer slips than Control when *EWR*=2 and *EWR*=3 (both $p < .001$), but that it was not significantly different from Control at *EWR*=1 ($p = 1.00$). Steady and Steadied-Bubble resulted in significantly fewer slips than control for all *EWR* (all $p < .005$), except at *EWR*=2, where Steady and Control

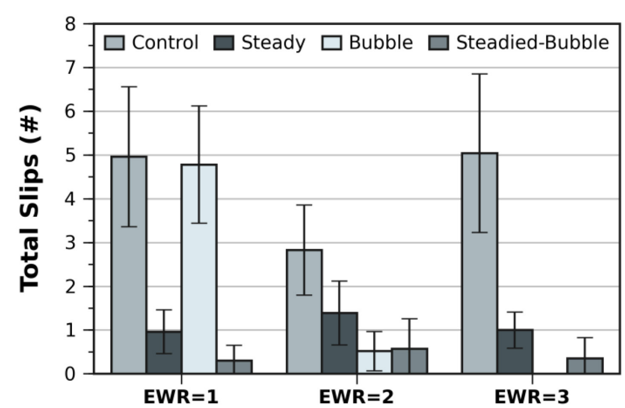

Figure 4. Average total slips (for 36 trials) by *CT* and *EWR* (*WID*=12, *N*=23). For *EWR>1*, all three experimental cursors performed comparably. AT *EWR*=1, Bubble was not effective.

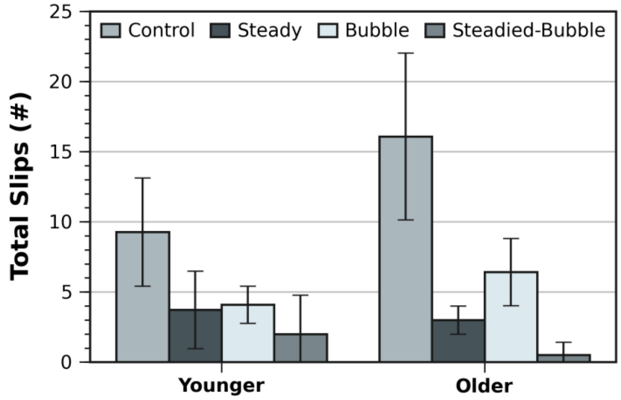

Figure 5. Average total slips (for 108 trials) by *CT* and *AGE* (*WID*=12, *N*=23). The only age difference was for Control.

were not significantly different (p = .14). Though this last comparison is somewhat inconsistent with our hypotheses, we note that as shown in Figure 4, it is mostly a reflection of lower than expected slip results for Control at *EWR*=2.

The older adults benefited more from the experimental cursors. As shown in Figure 5, there was a significant *CT* × *AGE* interaction ($F_{1.4,29.8}$ = 4.20, p = .036, η^2 = .167). Pairwise comparisons revealed that the experimental cursors reduced the performance gap between ages: For Control, the older group slipped significantly more than the younger group (p = .05), but there were no significant differences between the groups for any of the other interfaces. (A trend suggests a difference between the two groups for Bubble, p = .08. This likely reflects the influence of the Bubble cursor's increased slip rate at *EWR*=1.)

There was also a main effect of spacing (*EWR:* $F_{2,42}$ = 19.6, p < .001, η^2 = .482), indicating more slips at *EWR*=1. As with misses, this mirrors the main results.

Distracter Target Hits

Errors in Bubble and Steadied-Bubble were almost four times more likely to hit a distracter target. As shown in Figure 6, almost 100% of the errors in Bubble and Steadied-Bubble landed on a distracter target, while for Control and Steady the percentages were much lower: 28% and 23%, respectively. This difference was confirmed with an RM ANOVA on *AGE* and *CT* (main effect of *CT*: $F_{2.3,49.9}$ = 534, p < .001, η^2 = .960). Pairwise comparisons confirmed the percentages were higher for Bubble and Steadied-Bubble than for Control or Steady (all p < .001). (There were no main or interaction effects for *AGE*.)

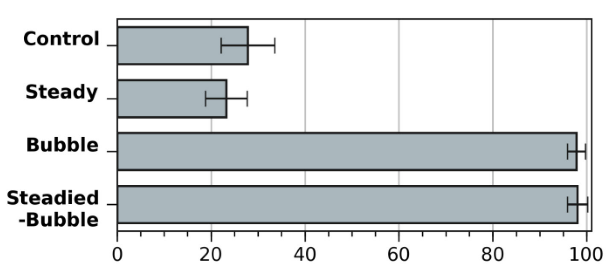

Figure 6. Average percentage of errors hitting a distracter target (*N*=23). Error bars are 95% Confidence Intervals.

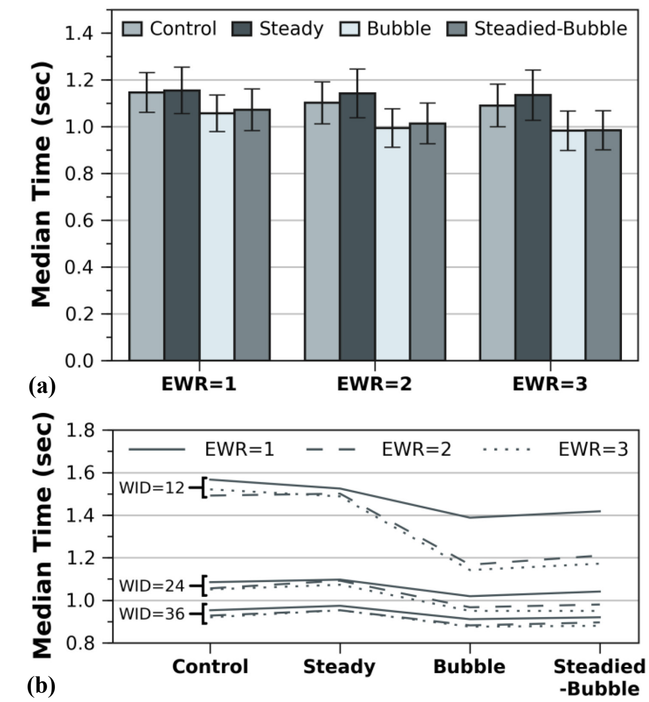

Figure 7. Average median trial time by (a) *CT* and *EWR,* and (b) *CT*, *EWR* and *WID*. Bubble and Steadied-Bubble were faster for *EWR*=2 and *EWR*=3. (*N*=24.)

Movement Time

Not including break times, the experimental tasks for each condition took on average 5.5 and 7.9 minutes, for the younger and older groups respectively (*SD* = 0.87, 1.25).

Consistent with the results for misses, Bubble and Steadied-Bubble were significantly faster than Control, except when targets were adjacent. There was a significant *CT* × *WID* × *EWR* interaction, as well as all the corresponding main and 2-way interactions (*CT* × *WID* × *EWR:* $F_{4.3,.81.1}$ = 7.27, p < .001, η^2 = .277; *CT*: $F_{3,57}$ = 31.9, p < .001, η^2 = .627; *WID:* $F_{1.0,19.4}$ = 291.5, p < .001, η^2 = .939; *EWR:* $F_{2,38}$ = 91.0, p < .001, η^2 = .827 *CT* × *WID*: $F_{2.1,39.8}$ = 20.1, p <001, η^2 = .514, *CT* × *EWR*: $F_{3.6,67.6}$ = 10.1, p < .001, η^2 = .346, *WID* × *EWR*: $F_{1.9,35.6}$ = 45.5, p < .001, η^2 = .706).

As shown in Figure 7a, cursors based on the bubble cursor were faster than Control and Steady, for *EWR*=2 and *EWR*=3. Figure 7b shows this same data broken down by *WID*, highlighting that the gains were largest at *WID*=12. Pairwise comparisons confirmed these differences. For *EWR*=2 and *EWR*=3, Bubble was significantly faster than both Control and Steady (all *WID*, p < .05) and Steadied-Bubble was significantly faster than Steady (all *WID*, p < .05) and Control *WID*=12 and *WID*=24, p < .05). At *EWR*=1, none of the comparisons were significant.

The older adults were slower than the younger adults and disproportionately affected by the task factors. There was a main effect of *AGE* and several interactions involving *AGE*. In all the interactions both age groups showed similar patterns of results, but the effects were magnified for the older adults. For both groups, speed decreased as target width increased and amplitude decreased, but the older group was disproportionately slowed by smaller targets and larger amplitudes. Moreover, inspection of the significant *AGE* × *CT* × *WID* interaction revealed that at the smallest target size, the older adults benefited more than the younger adults from Bubble and Steadied-Bubble. (*AGE*: $F_{1,19}$ = 41.5, $p < .001$, η^2=.686; *AGE* × *CT*: $F_{3,57}$ = 3.87, $p = .014$, η^2 = .169; *AGE* × *WID*: $F_{1.0,19.4}$ = 30.0, $p < .001$, η^2 = .612; *AGE* × *AMP*: $F_{2,38}$ = 4.22, $p = .022$, η^2 = .182, *AGE* × *CT* × *WID*: $F_{2.1,39.8}$ = 6.47, $p = .003$, η^2 = .254, *AGE* × *WID* × *EWR*: $F_{1.9,35.6}$ = 4.89, $p = .015$, η^2 = .205).

Finally, there was also a main effect of *AMP* ($F_{2,38}$ = 65.0, $p < .001$, η^2 = .774). Pairwise tests confirmed it was typical: Speed decreased as amplitude increased (both $p < .001$).

Pressure

The older adults exerted significantly more pressure than the younger adults. Although not our main focus, we additionally compared how much pressure younger and older adults exerted during selections. An independent t-test revealed a significant effect of *AGE* on the average maximum trial pressure (t_{22} = -3.34, $p = .003$, η^2 = .336). On average, the older adults applied approximately 50% more pressure than the younger adults (younger: M = 384, SD = 153; older: M = 595, SD = 157).

Subjective Findings

Bubble and Steadied-Bubble were preferred overall. A reliability analysis confirmed high consistency among the different rankings (Cronbach's alpha = .912), so we collapsed them into a single score for brevity. A Friedman test on the transformed rankings showed a significant main effect of *CT* (χ^2_3 = 35.27, $p < .0001$). Pairwise comparisons using Wilcoxon Signed Rank tests and a Bonferroni adjustment further revealed that Bubble and Steadied-Bubble were both ranked more favorably than Steady and Control, with no differences between either Bubble and Steadied-Bubble or Steady and Control (Bubble–Control: z = -3.88, $p < .001$; Bubble–Steady: z = -3.85, $p < .001$; Steadied-Bubble–Control: z = -3.5, $p < .005$; Steadied-Bubble–Steady: z = -4.03, $p < .001$).

Summary

We summarize our results for *WID*=12 according to our hypotheses. Recall, we were unable to test our hypotheses for the other target widths.

H1. *Bubble will reduce both slips and misses, but only when surrounding targets are not directly adjacent.* **Supported**.

H2. *Independent of target spacing, Steady will reduce slips, but it will not affect misses*. **Mostly supported**. Steady resulted in significantly fewer slips than Control, except there was no statistical difference at *EWR*=2.

H3. *Steadied-Bubble will reduce slips when targets are directly adjacent, and both slips and misses when they are not.* **Supported**.

H4. *The experimental cursors (Bubble, Steady, and Steadied-Bubble) will reduce total errors for both age groups, but the impact will be larger for the older group as they will benefit from both slip and miss reduction*. **Supported**.

H5. *A greater proportion of errors in Bubble and Steadied-Bubble will land on a distracter target*. **Supported**.

DISCUSSION

This study established the individual benefits of Steady Clicks and Bubble cursor for pen-based pointing with younger and older adults, and furthermore, showed that the two techniques can be successfully combined to provide the benefits of each. Our Steadied-Bubble cursor reduced misses when targets were not directly adjacent, and slips independent of spacing. Though our error analysis is limited to the smallest target size examined, we did see similar patterns for the larger target sizes. Moreover, our analysis of the movement time data found differences for all target widths. The pattern of results for movement time was similar to the one observed for missing. The results for slipping were not reflected in the movement times, which is not surprising since missing was the dominant error type.

Though the experimental techniques were beneficial to both age groups, they especially helped the older adults. For slipping, they worked so well that they reduced the performance gap between ages such that the older group was no longer significantly different from the younger group. For missing, both groups benefitted equally from the experimental cursors, but the older group missed almost three times more often. Thus techniques that reduce missing should have greater practical significance for them. We note this finding was not predicted, and it contrasts the results of our prior study that did not find an effect of age for misses [18]. One difference is that in [18], the smallest target was 38% bigger than in the current task.

Although slips were generally less frequent than misses (for both groups), slipping presents an important problem for older adults. During a slip, the pen initially lands on the target. This activates the visual feedback associated with a selection, and indicates to the user that their selection should be successful. As a result, slip errors are particularly confusing. Many older users are unaware of the cause of their difficulty, hindering self-correction strategies.

Thus, both slip and miss reductions are important for older users. The techniques we evaluated in this paper, and particularly the Steadied-Bubble addressed these two most common pen-based pointing problems. Importantly, none of our techniques hindered the younger participants (they were either positively or neutrally affected by all our cursors). Thus, inclusion of these techniques should make it easier for older adults to interact with the same software as younger adults, reducing the need for specialized software. Specialized programs generally try to make interaction easier by making targets bigger, often at the expense of aesthetics or features. However, they require each individual program to be adapted. Thus one major benefit of a pointing technique approach is that it provides older adults with access to a much broader set of applications.

Our finding that the older adults exerted 50% more force than the younger adults is important. Older adults often report finding pen interfaces tiring [5, 11, 18]. Older adults are known to have lower maximum force capabilities [13], and at first glance, this seems to explain why they might find them more tiring. However, our results suggest that the problem is not with exerting sufficient pressure, but rather with determining how much pressure is needed. In this study, we included visual feedback indicating contact. This type of feedback is not common; thus, our results may even underestimate the extent of the problem. Devising ways of teaching older adults to use less pressure is thus an important area for further investigation.

An additional finding was that although the two techniques based on the Bubble cursor reduced errors relative to the Control cursor, when errors did occur, they were almost four times more likely to result in selection of a distracter target. In contrast, a comparable effect was not observed for the Steady cursor. This tendency for the Bubble cursor to shift errors onto unwanted functionality has not been discussed in the literature to date, and it has important practical implications as selection of an unwanted target typically requires corrective action, and thus has a much higher cost than selection of inactive whitespace. This is particularly important for older adults as they tend to find error correction more difficult. The impact of Bubble and Steadied-Bubble's higher proportion of distracter target errors is likely not reflected in our preference ratings. Though we differentiated between hits on a distracter target and hits on inactive whitespace in our analysis, from the user's perspective these errors were the same. This may have contributed to the strong preference for the Bubble and Steadied-Bubble cursors over Steady and Control.

One place where our hypotheses were not fully met is at the medium level of target spacing examined (EWR=2), where Steady did not result in significantly fewer slips than Control. Inspection of the means for Steady and Control suggests it arose from Control performing slightly better at that spacing. Steady's performance remained relatively constant across spacings. Though the differences between levels of spacing for Control were not significant (and thus random variation is the likely explanation), it is possible that EWR=2 represents a balance between visual complexity at EWR=1 and overconfidence at EWR=3.

It is interesting that in our study we did not see a relationship between overall target density and performance for the Bubble cursor, while Grossman and Balakrishnan reported a negative effect of low density on performance [10]. One possible explanation is that limiting the maximum size of the Bubble cursor was effective as hypothesized by Grossman and Balakrishnan [10]. However, it is also possible this difference reflects a deeper distinction between mouse and pen interaction. With a pen, users can remain above the detectable range of the screen until late in the interaction. While the pen is out-of-range, the cursor was hidden in all conditions; thus intermediate distracter targets did not have as much of an effect on the behavior of the bubble cursor.

Finally, though we focused on Bubble cursor and Steady Clicks, many other mouse techniques have been developed, including target expansion [6, 16], Sticky Icons [26], and Object Pointing [9] (for an overview see [3]). Some of these may also have applicability to pen interaction, and there may be additional opportunities to combine them, as we have done here for Steady Clicks and Bubble cursor.

DESIGN CONSIDERATIONS

Our results show clear support for the experimental cursors and illustrate the particular task contexts in which it performed best. Thus, we conclude by reflecting on our findings to propose design guidelines for cursor selection.

Target Size and Density

We found that the biggest benefits were realized when targets were small; specifically, when they were comparable to the height of a text link. Even for the next biggest size examined (roughly the size of a small toolbar icon) the differences were weak and difficult to interpret. However, small targets abound and facilitating their selection is important.

To be effective, the Bubble cursor requires that targets are not directly adjacent. Thus, techniques based on it are particularly well suited to applications or tasks that have many small, but spaced targets. Such applications include interacting with large data visualizations, or selecting features in a drawing application. However, small targets are often coupled with high target density, such as with word or character selection in a text editor. In these situations, the Bubble cursor is not helpful, but techniques such as Steady Clicks are. The Steadied-Bubble provides a useful balance between these factors. For example, some web pages have links tightly clustered in one area, whereas other pages have sparser links. The Steadied-Bubble supports a seamless transition between these cases, providing the best possible support for each.

Error Cost

Overall, the Bubble and Steadied-Bubble cursors were more effective than the Steady cursor. However, because the Bubble cursor assigns inactive whitespace to nearby targets, it results in a higher proportion of errors landing on an unwanted target. In some cases, unintended selections are easily and efficiently corrected. However, it is important to consider the cost of error correction when choosing a technique. When the cost is high, it may be better to choose a technique that has more but less costly errors.

Target Awareness

As noted earlier, an important limitation of the Bubble cursor is that it must be target-aware; that is, it needs to know where targets are to expand the cursor. The Steady cursor is instead target-agnostic since it functions independent of target location. Thus, technique selection also depends on whether or not it is practically possible or computationally feasible to track target locations.

Target Users

The experimental cursors reduced the errors types we examined to differing degrees. In particular, Steady only provided support for slipping, and even with the Control cursor, the younger adults demonstrated relatively little slipping. As a result, interfaces targeted exclusively to younger users may not warrant a Steady or Steadied-Bubble approach. However, none of the techniques hindered the younger participants, and both slip and miss reductions are important for older users. Thus, when targeting older individuals—or a range of users—techniques that address both error types should be adopted.

CONCLUSION

Older adults form a growing demographic of computer users. Pen-based interfaces are an appealing platform for this group, but despite a multitude of advantages, many older individuals find pen interaction challenging. Error reduction is important because errors can carry a high cost for recovery and are overly frustrating for some users. We found that by combining existing techniques designed for younger users and mouse interaction, we substantially reduced the two most common types of pen-based error over a range of task contexts. We then drew upon these findings to identify guidelines for technique selection. These guidelines may also serve as a starting point for additional research in this area.

ACKNOWLEDGEMENTS

We thank L. Findlater, E. Murphy-Hill and T. Tang for comments on the paper and NSERC and CIHR for funding.

REFERENCES

1. Accot, J. & Zhai, S. 2002. More than dotting the i's—Foundations for crossing-based interfaces. *Proc. CHI'02*, 73–80.
2. Apitz, G. & Guimbretière, F. 2004. CrossY: A crossing-based drawing application. *Proc. UIST'04,* 312–12.
3. Balakrishnan, R. 2004. "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *Int'l J. Human Computer Studies,* 61(6): 857–874.
4. Birdi, K.S. & Zapf, D. 1997. Age differences in reactions to errors in computer-based work. *Behav. Inf. Technol. 16*(6): 309–319.
5. Charness, N., Holley, P., Feddon, J. & Jastrzembski, T. 2004. Light pen use and practice minimize age and hand performance differences in pointing tasks. *Hum. Factors, 46*(3): 373–384.
6. Cockburn, A. & Firth, A. 2003. Improving the acquisition of small targets. In *Proc. HCI'03*, 181–196.
7. Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edition, Lawrence Erlbaum.
8. Greenstein, J.L. 1997. Pointing devices. In *Handbook of human-computer interaction (1317–1348)*. Elsevier, Amsterdam.
9. Guiard, Y., Blanch, R., & Beaudouin-Lafon, M. *2004. Object pointing: a complement* to *bitmap po*inting in GUIs. *Proc. GI'04*, 9–16.
10. Grossman, T. & Balakrishnan, R. 2005. The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. *Proc. CHI'05*, 281–290.
11. Hourcade, J.P. & Berkel, T. 2008. Simple pen interaction performance of young and older adults using handheld computers. *Interact. Comput., 20*(1):166–183.
12. Keates, S. & Trewin, S. 2005. Effect of age and Parkinson's disease on cursor positioning using a mouse. *Proc. ASSETS'05*, 68–75.
13. Ketcham, C. & Stelmach, G. 2004. Movement control in the older adult. In *Technology for Adaptive Aging (64–92)*, National Academies Press, Washington DC.
14. Lee, M. & Dey, A. 2007. Providing good memory cues for people with episodic memory impairment. *Proc. ASSETS'07*, 131–138.
15. Massimi, M., Baecker, R. & Wu, M. 2007. Using participatory activities with seniors to critique, build, and evaluate mobile phones. *Proc. ASSETS'07*, 155–162.
16. McGuffin, M. J. & Balakrishnan, R. 2005. Fitts' law and expanding targets: Experimental studies and designs for user interfaces. *TOCHI, 12*(4):388–422.
17. Mizobuchi, S. & Yasumura, M. 2004. Tapping vs circling selections on pen-based devices: Evidence for different performance-shaping factors. *Proc. CHI'04*, 607–614.
18. Moffatt, K. & McGrenere, J. 2007. Slipping and drifting: Using older users to uncover pen-based target acquisition difficulties. *Proc. ASSETS'07*, 11–18.
19. Mynatt, E.D., Essa, I. & Rogers, W. 2000. Increasing the opportunities for aging in place. *Proc. CUU'00*, 65–71.
20. Ren, X. & Moriya, S. 2000. Improving selection performance on pen-based systems: A study of pen-based interaction for selection tasks. *TOCHI*, *7*(3):384–416.
21. Rowe, M., Lane, S. & Phipps, C. 2007. CareWatch: A home monitoring system for use in homes of persons with cognitive impairment. *Topics Geriatric Rehabil. Smart Technol., 23*(1): 3–8.
22. Smith, M.W., Sharit, J. & Czaja, S.J. 1999. Aging, motor control and the performance of computer mouse tasks. *Hum. Factors*, *41*(3): 589–596.
23. Trewin, S., Keates, S. & Moffatt, K. 2006. Developing Steady Clicks: A method of cursor assistance for people with motor impairments. *Proc. ASSETS'06*, 26–33.
24. Walker, N., Philbin, D.A. & Fisk, A.D. 1997. Age-related differences in movement control: Adjusting submovement structure to optimize performance. *J. Gerontol. Psychol. Sci. 52*(1): 40–52.
25. Wobbrock, J.O. & Gajos, K.Z. 2008. Goal crossing with mice and trackballs for people with motor impairments: performance, submovements, and design directions. *ACM TACCESS, 1*(4): 4:1–4:37.
26. Worden, A., Walker, N., Bharat, K. & Hudson, S. 1997. Making computers easier for older adults to use: area cursors and sticky icons. *Proc. CHI'97*, 266–271.
27. Yan, J.H. (2000). Effects of aging on linear and curvilinear aiming arm movements. *J. of Experimental Aging Research, 26*(4):393–408.