Hover or Tap?
Supporting Pen-based Menu Navigation for Older Adults

Karyn Moffatt, Sandra Yuen, 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 users, particularly older ones, still struggle with pen-based interaction. One type of error, *drifting*, occurs when users accidentally hover over an adjacent menu, causing their focus menu to close and the adjacent one to open. In this paper, we propose two approaches to address drifting. The first, *tap*, requires an explicit tap to switch menus, and thus, eliminates the possibility of a drift. The second, *glide*, uses a distance threshold to delay switching, and thereby reduce the likelihood of a drift. We performed a comparative evaluation of our approaches with a control interface. Tap was effective at reducing drifts for both groups, but it was only popular among older users. Glide surprisingly did not show any performance improvement. Additional research is needed to determine if the negative findings for glide are a result of the particular threshold used, or reflect a fundamental flaw in the glide approach.

Categories and Subject Descriptors
H.5.2 [**Information interfaces and presentation**]: User Interfaces — *Input devices and strategies.*

General Terms
Design, Experimentation, Human Factors.

Keywords
Target Acquisition, Pen-based Interaction, Tablet PC, Older Users, Inclusive Design, Universal Usability.

1. INTRODUCTION
Technology is increasingly being promoted as a means of addressing cognitive and sensory age-related impairments and enabling aging individuals to live more independently (e.g., [13, 21, 24, 26, 32]). Because they are small, mobile and powerful, pen-based technologies such as Personal Digital Assistants (PDAs) and Tablet PCs are appealing platforms for these endeavors. Moreover, direct pen-based input takes full advantage of hand-eye coordination and offers a familiar form of interaction [11]. When compared to a mouse, pen input has been shown to be particularly beneficial for older adults [5, 6].

Unfortunately, many older adults do encounter difficulties when using pen input [23]. This observation has motivated us to gain a better understanding of the challenges inherent to pen interaction, and to thereby build improved pen-based interfaces. In a previous study [23], we identified the types of difficulties some users, particularly older ones, experience with inductive pen-based input (such as with a Tablet PC). In this paper, we seek to address one of those difficulties, namely, *drifting*.

Drifting is a menu-navigation difficulty, involving the unintended invocation of a menu adjacent to the one in focus. This difficulty occurs as a result of the tracking capabilities provided by inductive pens. In contrast to passive pen technology (most often associated with PDAs), inductive pen technology can sense the location of the pen both when it is touching the screen's surface and when it is hovering in near proximity. When in this hovering (or tracking) state [3], pen technology behaves like a mouse when no mouse buttons are pressed; for example, the cursor's location is tracked and tool tips are dynamically displayed.

One notable use of the tracking state is to support *hover-switching* between menus. That is, when a menu is open, moving the cursor over any other menu head causes the currently open menu to close and the menu under the cursor to open, as shown in Figure 1. This can also be done with the mouse in the tracking state, and generally provides a quick and efficient way to scan through

Figure 1. When a menu is open, hovering the pen over a new menu causes the currently open menu (i.e., File) to close, and the menu under the cursor (i.e., Edit) to open. Drifting is the accidental invocation of this feature.

menus. However, with the pen, users tend to accidentally cursor (or drift) over the adjacent menu closest to their hand, causing the desired menu to close, and the adjacent one to open. Individuals tend not to expect the system to respond when they are not touching the screen. Thus, although the switching interaction is consistent between mouse and pen, it feels less natural with a pen.

Occlusion further impedes the interaction and contributes to drifting. When a user's hand occludes the menu contents, it is common for the user to lift the pen and move it away to read the contents, thus increasing the likelihood of a drift occurring. Moreover, because occlusion can obscure the action, it can be difficult for users to learn the cause of drifting. We believe this limits their ability to self-correct the problem.

An additional difference between mouse and pen is that the Microsoft Windows XP Tablet PC Edition provides automatic layout modifications based on the handedness of the user. The limitations of this layout protocol make it difficult to recognize when a drift has occurred, leading to further confusion. In particular, the alignment of menus is adjusted to minimize hand occlusion; thus, for a right-handed user, each menu is moved to the left, and aligned, if possible, with the right edge of the menu head (by default a traditional mouse-based interface left-aligns the menu with its menu head). However, menus can only be moved as far as the left edge of the screen. As a result, the leftmost menus tend to be in the same location (see Figure 2 for an example), reducing the visual shift that might alert the user to a drift.

In our previous study [23], 35 out of 36 of our participants drifted at least once during the study, and this behavior did not improve over the course of the study. Moreover, although drifting did not have an impact on the overall task accuracy, it did impede performance. Trial time almost doubled when drifting occurred. Finally, although drifting did affect everyone, older participants (those aged 55 and over) were disproportionately affected: they drifted significantly more often than the younger participants.

In this paper, we present two interfaces to address drifting: tap and glide. *Tap* deactivates the hover-switching functionality and requires an explicit tap to switch menus, thus eliminating the possibility of a drift. *Glide,* uses a distance threshold to delay switching, thereby reducing the likelihood of a drift. We performed a comparative evaluation of our approaches against a control condition for both younger and older users. The tap interface was effective at reducing drifts for both age groups, but it was only popular among older users. The glide interface, however, did not show any performance improvement. Additional research is needed to determine whether this is because using a threshold is the wrong approach, or because the particular threshold used was sub-optimal.

2. 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 the use of the hover space for interaction. Finally, we cover research on menu navigation.

2.1 Effects of Aging on Targeting Ability

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 concerning the speed-accuracy tradeoffs involved in movement control. They tend to be more conservative, and make more corrective sub-movements [31]. Older adults have also been found to cover less distance with their primary movement [18], to make many more sub-movements en route [17], to make less smooth movements [33], and to have difficulty staying on the target while clicking [28]. In addition, slower selection speeds have been attributed to lower peak velocities [17, 18], longer deceleration phases [18], and more pauses while homing in on the target [17].

To date, very little work has examined the use of pen-based systems with older adults. Charness et al. performed an age-related comparison of the mouse and the light-pen [5, 6], and found that the pen out-performed the mouse for all ages and that it 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, a position that is unnatural for many users. Modern Tablet PCs are designed to be more comfortable, and thus, should result in higher satisfaction. Additionally, Hourcade and Berkel [15] compared two pen-based selection techniques, tapping and touching (selection if the pen touches the target at any time before tap up), across age groups, and found that for the smallest target size examined (3.8 mm) the oldest group was more accurate using touch.

2.2 Hover-Space Interaction

Recently, there has been a growing interest in hover-space interaction as a means of increasing the functionality of pen-based systems [10, 12, 14]. For example, with Hover Widgets, users perform gestures above the surface of the screen to invoke functionality without moving away from their current work area [12]. Tracking Menus use the tracking state (of either a pen or a mouse) to enable a menu, such as a tool palette, to track the input device and remain close the user's work area [10]. One use of Stitching Gestures is to allow users to drag items (such as documents) between tablets with gestures in the hover-space [14]. Some research has also extended the capabilities of the hover space by exploring multiple layers of interaction [16, 30].

Research has also investigated the use of the hover space as a means for improving general target acquisition. For example, one investigation compared six target selection strategies, two of which used the hover-space [25]. However, their results did not show any clear benefit for the hover-space techniques.

2.3 Menu Navigation

Many researchers have investigated methods for improving menu interaction. Most of this research has focused on improving the efficiency of making a selection within a menu. For example, Split Menus reduce the average targeting time by placing the most frequently used items at the top of the menu [27], while Pie Menus decrease the distance the user must travel by arranging items circularly around the cursor [4]. Morphing Menus progressively grow the target area of frequently selected menu items over time, allowing them to borrow motor-space from neighboring infrequently used items [8].

When research has investigated how users navigate menus, it has focused on improving sub-menu navigation [1, 2, 7, 19]. This problem is quite different from ours. The challenge of sub-menu interaction is facilitating the invocation of desired sub-menus. Many users have difficulty remaining within the activation area necessary to complete this action. In contrast, we are looking for ways to prevent the accidental invocation of undesired menus. Nonetheless, looking at the characteristics of sub-menu interaction may provide insight into solutions for drifting.

3. PROPOSED SOLUTIONS

We have developed two approaches that we predict will address drift. First, we note that in our previous study [23], no one intentionally used hovering to switch between menus. Thus the simplest way to prevent drifting may be to turn off hover-switching, and require an explicit tap to switch between menus. This approach, which we call *tap*, clearly eliminates the drifting difficulty. However, the cost of this approach is not entirely clear. In our previous study, participants were all novices to pen-based interaction and were prompted to the correct menu (i.e., the task prompt for each trial provided both the menu and the item to be selected). It is possible that for expert users, or when the task requires browsing through menus for the correct item, being able to switch menus without touching the screen may prove useful.

An alternative approach, which we call *glide*, uses a distance threshold to differentiate between accidental drifts and intentional hover-switches. In our previous study [23], most drifts were short relative to the width of the menu head: over 80% of those drifts were fewer than 10 pixels (2.4 mm) into the adjacent menu head before the pen exited either the hover region of the tablet, or the top or bottom of the menu head. We predict that when intentionally switching to a new menu, (right-handed) users will bring the pen clear across the menu towards the right edge, so as to minimize hand occlusion and enable them to read the menu contents. Based on these two factors, we chose a threshold of 20 pixels (i.e., 4.8mm or approximately 40% of the width of the menu head). This is clearly larger than the majority of drifts in the previous study, and yet less than half way across the menu head.

We note that a time threshold may also be possible. We suspect that when drifting, users spend less time over the menu head than when they are intentionally switching to that menu. So, essentially, we could delay the menu switch by time, rather than distance. However, a time threshold is likely to require user specific customization, whereas we believe a distance threshold depends more on the width of the menu head than on individual differences in motor behavior. Thus, we chose to first investigate the simpler option, a distance threshold.

4. EXPERIMENTAL METHODOLOGY

To assess the effectiveness of our proposed designs, we ran a controlled laboratory experiment to compare the tap and glide interfaces relative to each other and to a control interface.

4.1 Participants

Our main focus for this investigation was on the effectiveness of our designs for older adults. However, we also wanted to ensure that these designs did not have a corresponding negative impact on performance for younger users. Thus, we recruited 24 participants from two age groups:

- *Younger* (aged 19–25, mean 21; 3 male, 9 female), and
- *Older* (aged 65–85, mean 72; 4 male, 8 female).

Participants received $5 for each ½ hour of participation. The younger participants were recruited through advertisements posted on campus, and completed the study in 60 to 90 minutes (mean: 75). The older participants were recruited though postings in the community and word-of-mouth advertisement. They took 100 to 150 minutes to complete the study (mean: 120). One older participant (not included in the 12 above) was unable to complete the study tasks. His data is not included in our analysis.

Participants were right-handed and free of diagnosed motor impairments to their right hand. 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 computer users and used a greater number of applications; older participants were more educated (in general) and had been using computers for longer. Surprisingly, there were no differences between the groups in terms of self-rated computer expertise.

4.2 Design

We used a repeated-measures design with interface condition (tap, glide, control) as a within-subjects factor. We chose this design for the increased power a within-subjects methodology provides and because it allows for comparative comments on the three interface conditions. We fully counterbalanced the presentation order of the interfaces to help control for learning effects.

Because younger participants are easier to recruit, we decided to design our study such that we could learn from the younger group, and if necessary further refine our study before focusing on the older adults. Thus, we ran the study as two independent experiments, fully completing the younger participants before starting with the older ones. Between the experiments, we reflected on the experimental design and made a small but important adjustment to how we referred to our interfaces.

We needed to provide an easy way to refer to our interfaces as we wanted participants to provide comparative feedback on all three. Initially, we named our interfaces tap, delay, and slide. We used *delay* to reflect the influence of the threshold on the menu-switching interaction. However, some participants in the younger group interpreted it to mean that the glide condition was inherently slower than the other conditions even though this was not supported by the performance data. Though this limits our ability to interpret the subjective measures for the younger group (particularly their response to the delay interface), we did not want it to also limit our analysis for the older group.

Thus, we renamed the *delay* condition to *glide*. This required we also rename the control condition from *slide* to *entry* to prevent confusion (between slide and glide), and to more clearly reflect the subtle distinction that in the control condition the menu changes as soon as you enter, whereas in the glide condition it happens at some point once you have entered. (Note, we did not specifically indicate to participants at what point the menu changes in the glide condition because we did not want users focusing on the threshold.) The following summarizes the original and revised names for each of our interfaces, and provides the short descriptions used to introduce participants to the conditions:

Tap (unchanged): The selected menu changes *when you tap* on a new menu.

Glide (was delay): The selected menu changes *as you move* the pen across a new menu, even while not touching the screen.

Entry (was slide): The selected menu changes *as soon as you move* the pen over a new menu, even while not touching the screen.

As the study was run in two distinct chronological phases, and as we modified the interface names between the phases, we are careful about performing age-related statistical comparisons in our analyses of the results (Specifically, in our main analyses, we perform separate statistical tests for each age group). However, establishing differences between younger and older users is not the main focus of this study. Age-related differences in drifting behavior are discussed in [23].

To summarize, our design consisted of two independent experiments (one per age group). Each experiment used a 3 (interfaces) x 6 (presentation orders) factorial design. Each participant was randomly assigned to one of the six possible presentation orders.

4.3 Procedure

The study began with a series of standardized tests of sensory-perceptual and motor-skills, and a questionnaire on background and computer experience. Participants were then introduced to the Tablet PC and completed steps 1–8 of "Get Going with the Tablet PC". (Steps 9–17 concern text input and were not relevant.) After the tutorial, the tablet was calibrated to each participant using the built-in Windows XP (Tablet PC Edition) calibration utility.

Participants then completed the menu conditions. At the start of each condition, a short description of that condition was provided (as described above in Section 4.2). Between conditions, participants completed short verbal distracter tasks. Finally, at the end of the study, participants were asked to rank the conditions, and encouraged to make additional comments. Beyond the instructions given in the tutorial, participants were not instructed to use the pen in any particular manner. We explicitly wanted to observe how they would naturally approach the task.

4.4 Task

The menu task was as follows. For each menu interface, participants completed a short 12-trial practice block followed by 6 blocks of trials with an enforced 45 second break between blocks. Each block consisted of a 36-item randomly ordered selection sequence from three 12-item menus. (Each item was selected once in each experimental block. The practice block consisted of a random subset of the items.) Thus, each participant completed 36 trials x 6 blocks x 3 interfaces for a total of 648 trials (excluding the 12 practice trials).

For each trial, a menu item was displayed across the top of the screen, as shown in Figure 2. Participants were instructed to find and select that item from the menus as quickly as possible while remaining accurate. The system advanced to the next item, only when the participant successfully selected the *correct* menu item. A soft clicking sound provided feedback for correct selections and a louder beep sound alerted participants to selection errors. Specifically, participants were not told which menu contained the target item. We wanted to ensure participants would need to search through the menus to find the correct item. This was done to encourage intentional use of hover-switching.

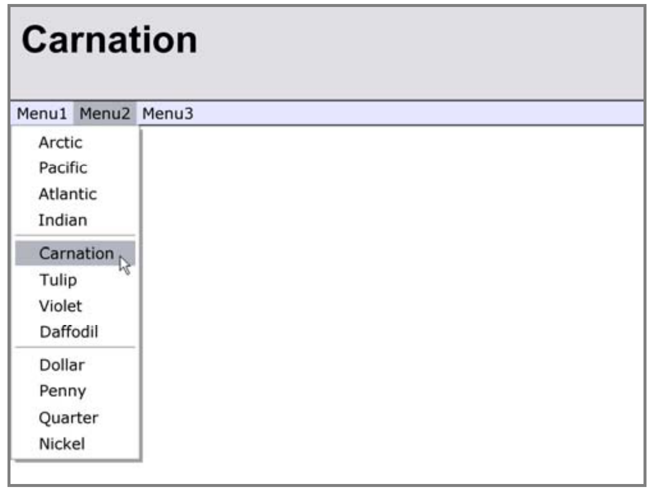

Figure 2. The upper-left corner of the study interface. Notice that Menu2 is aligned with the left edge of the screen; this is the default for a right-handed user on the Tablet PC.

Menu contents remained constant within each condition, but changed between conditions. Each menu contained three groups of four semantically related items. These schemes were randomly generated for each participant using the approach presented by Cockburn et al. [8]. That is, three 4-item groups were randomly selected from a collection of such groups. The items were then randomly ordered within that group, and no group was reused in any other condition. (See Figure 2 for an example of a generated menu.) Each menu item was 20 pixels (4.8 mm) high, and each menu separator was 5 pixels (1.2 mm) high.

4.5 Measures

Our main goal for this study was to examine the effect of each of our interfaces on drifting. As a measure of drifting, we recorded the number of extra target menu invocations for each trial; that is, the number of times the target menu was opened in excess of the once required to complete the task. Clearly, factors other than drifting can cause additional invocations. For example, the user may miss the target item on the first pass, or accidentally close the target menu before making a selection. However, we would not expect these factors to disproportionately affect any of our conditions; thus, we can interpret differences in the numbers of extra target menu invocations needed as an indication of differences in drifting behavior.

Additionally, we wanted to ensure we captured any unforeseen effects of our designs on other aspects of performance. Thus, we also included time as a dependent measure. An implicit error penalty was included by forcing participants to correctly select the target item before continuing to the next trial. Errors were also recorded independently for completeness. We note that in our previous study [23], drifting was found to significantly impact trial time. Thus, time may also be a useful indication of drifting difficulty. However, as drifting affects a relatively small number of trials and many other factors can affect trial time, we do not expect that overall task time will necessarily be sufficiently sensitive to detect difference between our conditions.

A poll-style questionnaire was administered after all conditions had been completed. Participants were asked to rank order each interface according to the following criteria: overall preference, speed, accuracy, frustration, and initial ease of use.

4.6 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 of performance was presented during the breaks between blocks.

4.7 Apparatus

All experimental conditions were run on a Fujitsu LifeBook T3010D Tablet PC with a 1.4 GHz Pentium M processor and 768 MB RAM, running the Windows XP Tablet PC Edition operating system. The display was 12.1 inches large, with a resolution of 1024 x 768. We removed the button on the side of the pen to ensure participants did not accidentally use it, as it was not required for the study tasks. The experimental software was written in Java, using the Standard Widget Toolkit (SWT).

The Tablet PC was placed on a stand, which positioned the screen at a comfortable viewing angle (based on previous pilot studies) of approximately 35 degrees from horizontal. We chose this setup, as it is difficult to view the screen when it is flat, and holding the tablet can be strenuous for older participants. Participants were encouraged to adjust the position of their chair and the stand for comfort; most made these adjustments.

4.8 Hypotheses

We had three main hypotheses for this experiment:

H1: Target Menu Invocations. We expect that both tap and glide will reduce drifting as measured by the number of extra target menu invocations.

H2: Speed and Accuracy. For speed, we predict that: (1) tap will be at least as fast as control as the cost of having to explicitly tap should be offset by reduced drifting, and (2) glide will be faster than control as it will reduce drifting without impeding hover-switching. For accuracy, we predict that there will be no differences among the interfaces.

H3: Subjective Response. We predict that both tap and glide will be preferred over the control interface. Between glide and tap, we predict there may be some age-related difference, with the older users tending to prefer the control provided by the tap interface, and the younger users preferring the efficiency of glide.

5. RESULTS

For each of our performance measures, we performed separate two-way repeated measures ANOVAs (Presentation order x Interface) for *each* age-group. Bonferroni corrections were used for all post-hoc pair-wise comparisons. Along with statistical significance, we report partial eta-squared (η^2), a measure of effect size, which is often more informative than statistical significance in applied human-computer interaction research [20]. To interpret this value, .01 is a small effect size, .06 is medium, and .14 is large [9].

Initial analysis of the data revealed a practice effect. We thus examined the data for differences between the blocks. We found that for the older group the first two blocks were significantly slower than the latter four, but that there were no significant differences among the last four blocks. For the younger group, only the first block was significantly slower than the others. Thus, in all subsequent analyses, we exclude the two blocks of each interface for the older group, and the first block for the younger group. This does not entirely eliminate the practice effect (i.e., there is still an interaction between interface and order on time, see Section 5.2); however, it reduces its impact.

5.1 Target Menu Invocations

Two-way repeated-measures ANOVAs (Presentation order x Interface) revealed a main effect of interface on the number of extra target menu invocations for both the older ($F_{2,12} = 12.0$, $p = 0.001$, $\eta^2 = 0.667$) and younger groups ($F_{2,12} = 5.94$, $p = 0.016$, $\eta^2 = 0.498$). As expected there was no interaction effect between interface and presentation order for either group. Post-hoc pair-wise comparisons further revealed that, for the older group, tap resulted in significantly fewer extra target menu invocations than control ($p = 0.013$). For the younger group, there was a similar trend between tap and control ($p = 0.079$).

Figure 3 shows the mean number of extra target menu invocations required per trial by interface and age group. The smaller variability observed in the tap condition suggests that much of the variance in the other two conditions is a result of individual differences in drifting behavior. The figure also shows that the older participants drifted more than the younger participants (as we would expect based on our previous findings [23]) and that the tap interface reduced this performance gap.

Thus, hypothesis H1 was supported for tap, but not for glide. Consistent with our predictions, tap reduced the number of unnecessary target menu invocations required, for both age groups. However, contrary to our predictions, glide did not provide any significant improvement for either group.

5.2 Speed and Accuracy

For the older group, a two-way repeated-measures ANOVA (Presentation order x Interface) on time revealed a main trend for interface ($F_{2,12} = 3.44$, $p = 0.066$, $\eta^2 = 0.364$) and a significant interaction between interface and order ($F_{10,12} = 6.55$, $p = 0.002$, $\eta^2 = 0.845$). For the younger group, the same analysis yielded no

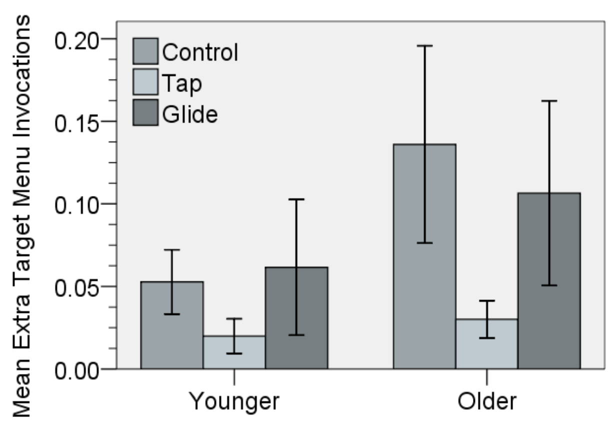

Figure 3. Mean number of extra target menu invocations required per trial by interface and age group (*N*=24). Error bars represent 95% confidence intervals.

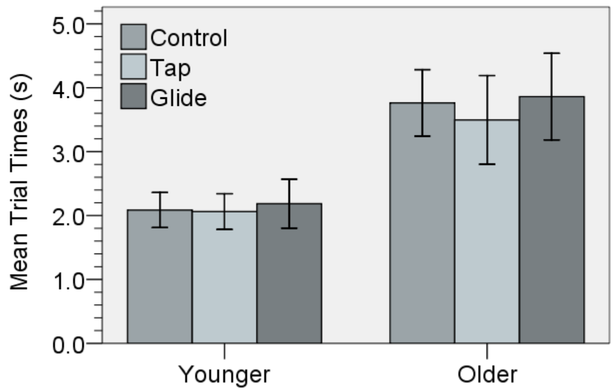

Figure 4. Mean trial time (s) by interface and age group (*N*=24). Error bars represent 95% confidence intervals.

significant results. Although accuracy is implicitly captured by speed, for completeness we examined the number of times an incorrect menu item was selected. As expected, there were no significant differences.

Figure 4 shows the mean trial times by interface and age group. Though it clearly highlights the large variability in this measure, it does support a trend towards the tap interface being faster for the older age group. Figure 5 shows mean trial times for the older group by interface and presentation order, and more clearly demonstrates the interaction effect observed. To summarize, tap was faster than control in all three orders where control was presented before tap, but control was faster than tap only in the condition where tap was presented first and control last, suggesting that practice was playing a role in the overall speed improvements. Glide performed comparably to control in all orderings except glide-control-tap, where it was slower than both tap and control. With only 2 individuals per order and age group; individual differences likely played a role here.

Thus, hypothesis H2 was also supported for tap, but not for glide. Consistent with our predictions in hypothesis H2.1, tap was not slower than control. Moreover, a trend suggests there may in fact be a small speed benefit for the tap interface for the older group. However, further research is necessary to confirm this trend, as there was also a significant interaction between interface and order. In contrast to our predictions in hypothesis H2.2 (but

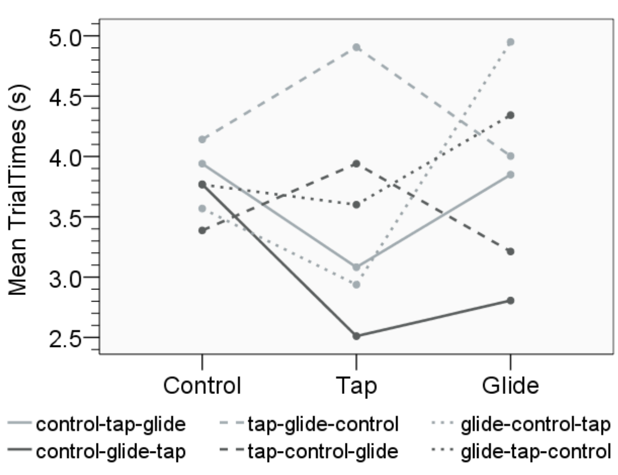

Figure 5. Mean trial time (s) by interface and order (*N*=24).

consistent with our negative findings for hypothesis H1) the glide interface was not faster than either of the other two interfaces. As expected there were no differences for accuracy.

5.3 Age-related Differences

Although age-related differences are not our main focus, for completeness, we re-examined our variables of interest with age as between-subjects factor. Specifically, we performed a three-way repeated-measures ANOVA (Age group x Presentation order x Interface) on the time and extra target menu invocations required. Not surprisingly, there was a main effect of age on both measures (time: $F_{1,12} = 31.7$, $p < 0.001$, $\eta^2 = 0.726$; invocations: $F_{1,12} = 8.34$, $p = 0.014$, $\eta^2 = 0.410$). For extra target menu invocations, there was also an interaction between age and interface ($F_{2,24} = 4.30$, $p = 0.025$, $\eta^2 = 0.264$) suggesting that the tap interface disproportionately helped the older group. The remaining results were consistent with our separate analyses. These results should be considered preliminary due to the design limitations described in Section 4.2.

5.4 Subjective Response

After all three conditions, participants were asked to rank the interfaces according to five measures: preference, speed, accuracy, frustration, and initial ease. Table 1 summarizes their responses. We analyzed these results using the randomization test of goodness-of-fit [22, 29]. A chi-square test of goodness-of-fit is more commonly used for this type of analysis; however, our expected values (4) were too low for this test (which requires a minimum of 5). The randomization test uses Monte Carlo simulation to calculate the probability of the observed frequency values occurring by chance. It is robust against small sample sizes and low expected values [29].

Hypothesis H3 was partially supported. The tap interface was well received by the older group, but there was some evidence of a negative response to it by the younger group. The older group rated tap as the most preferred ($p = 0.009$), the fastest ($p = 0.010$), the least frustrating ($p = 0.003$), and the easiest initially ($p = 0.022$). In contrast, the younger group perceived tap as the slowest ($p = 0.037$), and though not significant, it was also the least preferred by a majority of younger participants (7/12).

Table 1. Summary of self-reported preferences[†] (*N*=24).

Dependent Variable	Number of Votes					
	Younger Group			*Older Group*		
	Control	Tap	Glide	Control	Tap	Glide
Most Preferred	6	3	3	0	8	3
Least Preferred	1	7	4	4	3	5
Fastest	7	1	4	1	8	1
Slowest	0	7	4	3	3	6
Most Accurate	5	2	4	1	6	2
Least Accurate	5	3	3	2	1	7
Most Frustrating	2	5	4	5	1	5
Least Frustrating	7	3	1	0	8	1
Easiest Initially	4	5	3	0	7	3
Hardest Initially	3	5	4	5	2	5

[†] Some participants did not answer all questions; thus some rows sum to less than 24. Significant rankings are denoted by a grey background.

6. LIMITATIONS

One important limitation of this study is the null performance result for the glide interface. One possible interpretation is that using a distance threshold is the wrong approach and it is not possible to distinguish between accidental drifts and intentional menu invocations by distance. However, we did only try one threshold. Although it was chosen based on our previous data, it is possible that a different threshold may have yielded better results. Future research could further explore different thresholds, perhaps incorporating elements of both time and distance.

A second limitation is the significant practice effect that hindered performance in participants' first session. In retrospect, it is clear that we needed to provide more training. However, there is always a tension between maximizing training and minimizing study length. For the older adults, our study was an average of two hours long. It is not clear we could have made it much longer.

Finally, we note that we have focused exclusively on novice and right-handed individuals. This should not be interpreted as an indication that drifting is unique to them. We have excluded experts and left-handed individuals from our studies because we have been unable to properly control for these factors, and we do believe that there are likely some differences. Additional research is necessary to determine whether and how our results generalize to these populations.

7. DISCUSSION

In this section, we discuss our findings, focusing on their implications for design and on avenues for future research.

7.1 Eliminating drifting difficulties

Our results suggest that pen-based menus can be designed to prevent drifting and accommodate the needs of older adults, without impeding browsing. The success of the tap interface for reducing excess target menu invocations for both age groups suggests that hover–switching could be removed entirely without compromising overall menu navigation.

However, the tap interface was not well received by the younger group: they ranked it slowest and least preferred. What is not clear is the magnitude of this preference; further research would be needed to determine whether this preference is mild or strong. In the meantime, a safer approach may be to support personalization and allow users to turn off hover-switching. One limitation of a personalization approach is that it is unclear whether individuals (and particularly older adults) can self assess drifting behavior. In both this study and in our previous work [23], users seemed largely unaware of why the wrong menu sometimes opened. Thus, one avenue for future investigation would be to examine whether drifting can be inferred from user input. If so, the system could preemptively present this information to the user and offer to deactivate hover-switching.

Finally, the contradictory preference results observed in this study are noteworthy as they underline the critical importance of evaluation across the lifespan. Although the tap interface benefited both groups, had the study only focused on the younger demographic, the negative preference rankings for that group would have made interpretation difficult. Including both groups allowed us to see that the tap interface does provide benefit, but likely the magnitude of that benefit is small enough for the younger group that personal preference dominates.

7.2 Accessible hover-space interaction

Although our work has focused on the negative potential of hover-space interaction, not all such interaction is necessarily detrimental to the performance of older adults. Many of the proposed uses of the hover space have been to provide shortcut-style access to features that can also be accessed via more standard mechanisms (like with a button) [10, 12, 14], thus users need not rely on interacting within the hover space. Moreover, some hover-space research has paid careful attention to designing hover commands that cannot easily be triggered by accident. For example, Hover Widgets were designed so that their invocation gestures were both sufficiently simple to perform and sufficiently unambiguous that they would not be activated unintentionally [12]. Although it is unlikely that these interfaces will be usable by older adults, the combination of unobtrusive design and alternate access mechanisms should ensure that, at a minimum, they do not hinder performance.

7.3 Generalization to other interfaces

Finally, although we limited our investigation to menu interaction, we note that there are similarities between menu interaction and other interfaces (e.g. toolbars). With pen-based interaction, the combination of small widgets in close proximity, hand occlusion of the task area, and ambiguity in the interaction, may all contribute to the drifting difficulty. Proximity of the widgets is an important factor: as the distance between widgets increases, small movement away from the target is less likely to result in a drifting. Similarly, if hand occlusion were not present or if users did not need to read the contents of the menu, then there would be less of a need to move away from the target area in the first place. Finally, hover-switching is ambiguous; drifting would not be a problem if the required action were not one the user was likely to perform unintentionally. Designers should consider all these factors when building pen-based interfaces to avoid introducing similar problems.

8. CONCLUSION

This paper presents the findings of an experiment comparing the effectiveness of two techniques designed to reduce accidental menu invocations caused by drifting; that is, unintentionally cursoring over an adjacent menu head while in the tracking state. The tap interface, which turned off hover-switching and required users to make an explicit tap to switch between menus, was effective for both younger and older users in terms of reducing extra target menu invocations (an indication of drifting difficulty). However, despite the performance benefit, younger users did not like the tap interface.

The glide interface, which used a distance threshold to differentiate between accidental drifts and intentional menu invocations, did not show a performance benefit. Further research is needed to examine whether this is because using a threshold is the wrong approach, or because the particular threshold used in this experiment was sub-optimal.

In conclusion, pen-based devices are increasingly used in the development of assistive technology for cognitive and sensory age-related impairments (e.g., [13, 21, 24, 26, 32]). However, for this to be a viable approach, the accessibility of these devices needs to be improved for older adults. In our work, we have shown that improvements can be made to these interfaces without hindering their efficiency for younger adults.

9. ACKNOWLEDGMENTS

This research was made possible by support from the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

10. REFERENCES

[1] Ahlström, D. 2005. Modeling and improving selection in cascading pull-down menus using Fitts' law, the steering law and force fields. In *Proc. CHI '05*, 61–70.

[2] Ahlström, D., Hitz, M., and Leitner, G. 2006. An evaluation of sticky and force enhanced targets in multi target situations. In *Proc. NordiCHI '06*, 58–67.

[3] Buxton, W. 1990. A three-state model of graphical input. In *Proc. Interact '90*, 449–456.

[4] Callahan, J., Hopkins, D., Weiser, M., and Shneiderman, B. 1988. An empirical comparison of pie vs. linear menus. *Proc. CHI '88*, 95–100.

[5] Charness, N. Bosman, E.A. and Elliott, R.G. 1995. Senior-Friendly Input Devices: Is the Pen Mightier than the Mouse? *Presented at the 103rd Annual Convention of the APA.*

[6] Charness, N., Holley, P., Feddon, J., and Jastrzembski, T. 2004. Light pen use and practice minimize age and hand performance differences in pointing tasks. In *Hum. Factors*, 46(3): 373–384.

[7] Cockburn, A. and Gin, A. 2006. Faster cascading menu selections with enlarged activation areas. In *Proc. Graphics Interface '06*, 65–71.

[8] Cockburn, A., Gutwin, C., and Greenberg, S. 2007. A predictive model of menu performance. In Proc. CHI '07, 627–636.

[9] Cohen, J. 1973. Eta-Squared and Partial Eta-Squared in Communication Science. In *Human Communication Research*, 28, Oxford Journals, 473-490.

[10] Fitzmaurice, G., Khan, A., Pieké, R., Buxton, B., and Kurtenbach, G. 2003. Tracking menus. In *Proc. UIST '03*, 71–79.

[11] Greenstein, J.L. 1997. Pointing devices. In M.G. Helander, T.K. Landauer, & P.V. Prabhu (Eds.), *Handbook of human-computer interaction* (pp 1317–1348). Elsevier, Amsterdam.

[12] Grossman, T., Hinckley, K., Baudisch, P., Agrawala, M., and Balakrishnan, R. 2006. Hover widgets: Using the tracking state to extend the capabilities of pen-operated devices. In *Proc. CHI '06*, 861–870.

[13] Hawkey, K., Inkpen, K.M., Rockwood, K., McAllister, M., and Slonim, J. 2005. Requirements gathering with Alzheimer's patients and caregivers. In *Proc. ASSETS '05*, 142–149.

[14] Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. 2004. Stitching: pen gestures that span multiple displays. In *Proc. AVI '04*, 23–31.

[15] Hourcade, J.P., and Berkel, T.R. 2008. Simple pen interaction performance of young and older adults using handheld computers. In *Interact. Comput.*, 20(1):166–183.

[16] Kattinakere R. S., Grossman, T., Subramanian, S., 2007. Modeling steering within above-the-surface interaction layers. In *Proc. CHI '07*, 317–326.

[17] Keates, S. and Trewin, S. 2005. Effect of age and Parkinson's disease on cursor positioning using a mouse. In *Proc. ASSETS '05*, 68–75.

[18] Ketcham, C., and Stelmach, G. 2004. Movement control in the older adult. In *Technology for Adaptive Aging*, Washington DC: National Academies Press, 64–92.

[19] Kobayashi, M. and Igarashi, T. (2003). Considering the direction of cursor movement for efficient traversal of cascading menus. In *Proc. UIST '03*, 91–94.

[20] Landauer, T.K. 1997. Behavioral Research Methods in Human-Computer Interaction. In M.G.Helander, T.K. Landauer, and P.V. Prabhu, (Eds), *Handbook of human computer interaction* (pp 203–227). Elsevier, Amsterdam.

[21] Lee, M. and Dey, A. 2007. Providing Good Memory Cues for People with Episodic Memory Impairment. In *Proc. ASSETS '07*, 131–138.

[22] Manly, B.F. J. 1997. Randomization, bootstrapping, and Monte Carlo methods in Biology. Chapman and Hall, NY.

[23] Moffatt, K., and McGrenere, J. 2007. Slipping and Drifting: Using Older Users to Uncover Pen-based Target Acquisition Difficulties. In *Proc. ASSETS '07*, 11–18.

[24] Moffatt, K., McGrenere, J., Purves, B., and Klawe, M. 2004. The participatory design of a sound and image enhanced daily planner for people with aphasia. In *Proc. CHI '04*, 407–414.

[25] Ren, X. and Moriya, S. 2000. Improving selection performance on pen-based systems: A study of pen-based interaction for selection tasks. In *ACM TOCHI*, 7(3):384–416.

[26] Rowe, M., Lane, S., and Phipps, C. 2007. CareWatch: A Home Monitoring System for Use in Homes of Persons With Cognitive Impairment. In *Topics in Geriatric Rehabilitation: Smart Technology*, 23(1): 3–8.

[27] Sears, A. and Shneiderman, B. 1994. Split menus: effectively using selection frequency to organize menus. In *ACM TOCHI*, 1(1): 27–51.

[28] Smith, M.W., Sharit, J. and Czaja, S.J. 1999. Aging, Motor Control and the Performance of Computer Mouse Tasks. *Hum. Factors*. 41(3): 589–596.

[29] Sokel, R.R. and Rohlf, F.J. 1995. Biometry: The principles and practice of statistics in biological research (pp 803–820). Freeman, New York.

[30] Subramanian, S., Aliakseyeu, D., and Lucero, A. 2006. Multi-layer interaction for digital tables. In *Proc. UIST '06*, 269-272.

[31] Walker, N., Philbin, D.A., and Fisk, A.D. 1997. Age-related differences in movement control: Adjusting submovement structure to optimize performance. In *J. Gerontol*. 52B(1): P40–52.

[32] Wu, M., Baecker, R., and Richards, B. 2007. Designing a cognitive aid for and with people who have anterograde amnesia. In *Universal Usability*, Jonathan Lazar, Ed. John Wiley & Sons, 317–356.

[33] Yan, J.H. 2000. Effects of Aging on Linear and Curvilinear Aiming Arm Movements. In *J. of Exp. Aging Research*, 26(4): 393–408.