Of mice and pen: effects of input device on different age groups performing goal-oriented tasks

Convertible tablet PCs can use a pen or a mouse for input. The pen is better suited than the mouse for some tasks because of its interaction properties, and research has shown it may ameliorate age-related decrements in performance. This study compared the pen and mouse on a series of realistic tasks for older (55-69) and younger (18-30) adults. Precision tasks were better served by the mouse, while ballistic tasks with strong analogs to real-world actions were served equally well by the pen or the mouse. Older adults were slower than younger adults on both devices, but contrary to the research hypothesis, no benefits were observed specifically for older adults with the pen. This study reinforces findings regarding the importance of task demands when selecting input devices. Younger adults seemed more willing than older adults to embrace the pen.

INTRODUCTION

An integral component of the graphic user interface (GUI) is an input device for manipulating it, and since the GUI's introduction to personal computing, the mouse has been the device of choice. Yet, despite its ubiquity, even from the early days of the personal computer, the mouse has been a poor tool for certain tasks – e.g. drawing (Meyer, 1995). An important question for computing today is what types of input devices best support users in performing the ever broadening spectrum of tasks for which personal computers are being used.

The introduction to the market of the convertible tablet PC provides an opportunity to look at a computer that affords both traditional mouse input as well as the more novel pen-based input. Because tablet PC with pen computing is relatively new, however, there are few published studies that look at the current device technology, and studies of similar devices, while informative, may not be directly applicable or up to date. However, properties shared between earlier technologies and tablet PC computing can provide a basis for deriving insights from earlier studies.

Direct and indirect interaction

Input device interaction may be characterized as either direct (where the input surface is also the output surface) or indirect (where the input surface is different from the output surface). Direct interaction is most like using an analog tool (e.g. pencil and paper, hammer and nail, mechanical buttons) and is therefore more natural to use for analog-type tasks (e.g. drawing) and ballistic tasks (e.g. pointing at targets). A tablet PC with pen is characterized by direct interaction, whereas a mouse is indirect.

Benefits of direct interaction include minimizing issues of eye-hand coordination, stimulus-response compatibility, and learnability (Greenstein, 1997). However, parallax problems, ergonomic issues associated with the physical demands of the interaction, and occlusion of part of the display by the hand and arm are often cited as problems with direct interaction (Bullinger, Kern, & Braun, 1997; Greenstein, 1997; Hinckley, 2003). These problems of direct interaction can make it difficult to use for precision work.

Older adults are often identified as potential beneficiaries of direct interaction devices because these devices are easier to understand and require no mapping or translation, cognitive functions that are known to be slower in older adults (Proctor, 2005; Vercruyssen, 1997). However, two recent elemental task studies comparing touch screens to a mouse for older users present starkly contrasting results. Phillips (2003) found that older adults not only performed better on a pointing task with the mouse, but preferred it as well. Murata and Iwase (2006), however, found that for older adults, pointing on a touch screen was faster than mousing.

Older adults and input devices

The factor of age is particularly important for several reasons: aging is inevitable and universal, and the time is quickly approaching when, for the first time, large numbers of people will be entering more advanced age groups and using computers. Every study that has compared older adults to younger adults in computer interaction has found that older adults perform more slowly (Walker, Millians, & Worden, 1996; Chaparro, Bohan, Fernandez, Choi, & Kattel, 1999; Smith, Sharit, & Czaja, 1999; Charness, Holley, Feddon, & Jastrzembski, 2004; Rogers, Fisk, McLaughlin, & Pak, 2005). Additionally, research on older adults and input devices suggests that older users may have more difficulty with the mouse than younger users (Smith et al., 1999), and that alternate input devices may mitigate these differences (Chaparro et al., 1999; Charness et al., 2004). When using a light pen as opposed to a mouse, for example, older adults' decrement in performance compared to younger adults was found to be significantly reduced (Charness et al., 2004). As the literature is equivocal on the benefits of direct interaction for older adults in general, yet selection of certain input devices has been demonstrated to benefit older adults when compared with younger adults, more research in these areas is warranted.

A goal-oriented approach in device studies

The majority of studies on input devices have measured movement time and errors in a barrage of repetitive elemental tasks, such as pointing, dragging, and clicking (e.g. Chaparro, 1999; Charness, 2004; Cohen, Meyer, & Nilsen, 1993;

MacKenzie, Sellen, & Buxton, 1991; Murata & Iwase, 2006; Phillips, 2003; Smith et al., 1999; Walker et al., 1996). While these elemental studies provide baseline information about human performance with input devices, they do not illuminate the user experience with devices, and have limited external validity.

The advantage of a more goal-oriented approach is that the results provide insight into how the variables under consideration would affect performance in a more realistic context. However, identifying such tasks that are meaningful to a broad range of domains can be difficult. In an exploratory study of work-related reading, Adler, Gujar, Harrison, O'Hara, & Sellen (1998) looked at fifteen participants from a broad range of professions and identified some commonalities across domains. Among their findings, they identified reading for cross-referencing as the most frequently occurring task. In their taxonomy, "cross-referencing" involved reading from multiple sources or from one source in order to integrate information. By focusing on these types of tasks, an investigator may be confident that the results will be meaningful to a broad cross-section of users.

The present study

This study looked at a substantially different alternative to the mouse – the pen with tablet PC – in a range of goal-oriented tasks reflecting both commonly performed discrete tasks as well as more complex cross-referencing tasks common to many domains of work. Different age groups were compared to further investigate the potential benefits of alternative input devices for older adults.

Across all the tasks, it was hypothesized that older adults would perform slower than younger adults. Depending on the task, it was further hypothesized that the magnitude of age effects will depend on the device being used, such that use of the pen would reduce age-related decrements in performance compared to the mouse.

Device effects were hypothesized to vary depending on the task demands. Tasks that require more complex, abstract, or precise activity were hypothesized to be better performed with the mouse because the mouse uses indirect interaction. The pen, on the other hand, was hypothesized to be superior for tasks that involve primarily simpler, ballistic movements, and for tasks that are closely analogous to their real-world counterpart, such as drawing or dragging.

METHOD

Design

Independent variables. The study employed a 2x2 mixed factorial design, with the within-subjects factor of input device (pen, mouse) and the between-subjects variable of age group (younger, older). Six goal-oriented tasks were performed. However, as the tasks incorporated various elements, they were analyzed separately, and task was not considered a separate factor.

Dependent variables. Performance data was collected in the form of time on task and errors in task execution. An error analysis was conducted for each task, specific to the task demands. After completing all tasks with both devices, participants' overall device preference for each task were recorded. Subjective ratings of usability dimensions were collected for each device, as well as ratings of perceived effort (RPE), however their import is beyond the scope of this article.

Participants

A convenience sample of 24 participants (12 aged 18-30 with mean age of 21.17, 12 aged 55-70, with mean age 62.17) were recruited for the study. Participants were screened for demographics, computer experience, and physical limitations, to assure a level of competence for performing two hours of computer work, and were also tested for cognitive deficiencies by a digit span test. The younger participants received college credit for participating in the study. Older participants were paid $50 for 120 minutes of time.

The device sequence was counterbalanced within experimental groups, resulting in two possible orders (pen first, mouse first). Participants within each group were randomly assigned to one of the orders with the constraint that there be equal number of participants assigned to each order.

Apparatus

A Toshiba Satellite R25-S3503 convertible tablet PC with a 14.1" display (1440x900 screen resolution) was used for both input device conditions. For the pen condition, the PC was configured as a tablet, and the pen that comes supplied with the PC was used for input. For the mouse condition, the PC was configured as a regular laptop computer, with a Kensington Mouse-In-A-Box optical two-button mouse.

Six goal-oriented tasks were performed using Windows XP Tablet Edition operating system, Microsoft Excel, Microsoft PowerPoint, Microsoft Digital Image, and Microsoft Word applications. The last 2 tasks required the participant to cross-reference a printed document with an electronic one, and make changes to the electronic document based on the printed version.

Overall device preference by task was determined by a forced-choice pencil-and-paper survey which required participants to choose a favorite device for each task.

Procedure

Upon arrival, participants were greeted and asked to sign a consent form to participate in research. Next, the digit span test was verbally administered. A brief description of the study was provided, followed by an explanation of data collection instruments.

Participants were then familiarized with the computer and both input devices so that they could demonstrate that they were comfortable using either device. For the tablet PC conditions, participants were encouraged to experiment with various positions of the device.

Participants were all experienced with the mouse and required little practice or training with it. However, for most participants, this was the first time they had used the pen or a tablet PC, so all participants were given an introduction to tablet computing tutorial that comes with the computer. Following the tutorial, they were given practice time with the pen until they indicated that they felt comfortable using the pen.

After the familiarization period, participants performed the series of goal-oriented tasks with the first device. Prior to starting each task, participants were given a simplified practice task that gave them practice on all of the interactions they would need for the actual task. A required 5 minute break followed use of the first device, then the second series of tasks were conducted using the second device. Following use of both devices, the final overall choice survey was administered, and a debriefing interview was conducted. Participants were thanked for their time, and paid, if appropriate.

Task Descriptions and Analysis

For each task a GOMS-type analysis was conducted for the purpose of gaining insights into the task demands, and predicting which tasks would be better served by which input device.

Excel dragging task. Select and drag numbered cells from one position to another. Because of the amount of precision required to select the correct tool, it was hypothesized that the mouse would be the better device for this task.

PowerPoint dragging task. Drag 29 shapes into a constrained space, so that none of them touch. Because the task involves primarily ballistic movements and dragging movements, it was hypothesized that the pen will be the better device for this task. An interaction was hypothesized between age and device.

PowerPoint manipulation task. Move "callout" shapes to targeted spots on the screen, and manipulate the "tail" of the shape. Because of the amount of precision required to select the correct tool and manipulate the callout shapes, it was hypothesized that the mouse would be the better device for this task.

Tracing task. Trace a line-art image with curves, straight lines and points. Because of the strong analog between this task and drawing with a pen, as well as the continuous nature of the task, it was hypothesized that the pen would be the better device for this task. An interaction was hypothesized between age and device.

Cross-referencing form filling. Entering data from a printed survey document into an electronic, web-based, equivalent form. It was hypothesized that younger adults, with their better motor control will benefit by using the pen on the primarily ballistic task. However, older adults, who may have more difficulty with the precision required to hit the small targets, were expected to perform equally well with either device.

Cross-referencing document editing. Make changes to an electronic word processing document by referring to a printed document that has been marked up (or annotated) for layout changes. Because of the amount of precision required to accomplish one of the annotations, it was hypothesized that the mouse would be the better device for this task.

RESULTS

Data preparation and error analyses

Each participant's session was reviewed for an error analysis. Errors were classified as slips, lapses, or mistakes (Reason, as cited in Stanton, 2003). Slips (errors in execution but not intent) were the error type of interest in this study, as they reflect a failure of the input device to support the user in achieving their goal. Some participants were removed from some tasks due to a high number of mistakes (lack of understanding of the task).

Performance measures

To test the research hypotheses, for each measure on each task, a 2x2x2 mixed-factor ANOVA was conducted with age group (younger, older) and order (pen first, mouse first) as the between-subjects factors, and device (pen, mouse) as the within-subjects factor. In cases where differential carryover effects were present (an interaction of order and device), the analysis was repeated, using only data from the first device used. If results from the secondary analysis were in conflict with the original analysis, the secondary results are presented here.

Timing. Older adults were significantly slower than younger adults on all but the tracing task, where there was no difference between age groups. The mouse was faster for Excel dragging and PowerPoint manipulation, while the pen was faster for tracing. None of the interactions were significant. Refer to Table 1.

Table 1. Results of ANOVA for timing by task.

Task	Statistic	Age	Device	Age x Device
Excel drag	F (1, 19)	17.18**	15.34**	1.67
PPT drag	F (1, 20)	16.06**	.59	1.57
PPT manip.	F (1, 17)	18.97***	12.05**	2.03
Tracing	F (1, 20)	.94	27.28***	.88
Form filling	F (1, 20)	28.97***	.67	1.38
Doc. edit	F (1, 18)	31.00***	3.52	.55

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 2. Results of ANOVA for error score by task.

Task	Statistic	Age	Device	Age x Device
Excel drag	F (1, 19)	4.50*	11.84**	.49
PPT drag	F (1, 20)	3.25	12.07**	.54
PPT manip.	F (1, 17)	.45	9.08**	.06
Tracing	F (1, 20)	4.68*	63.02***	.67
Form filling	F (1, 20)	2.49	4.30	2.06
Doc. edit	F (1, 18)	3.93	12.30**	3.43

* $p < .05$ ** $p < .01$ *** $p < .001$

Errors. For all but two of the tasks, there were no significant error differences between age groups. In both the Excel dragging task and the Tracing task older adults made more errors than younger adults. However, significant device effects were observed in almost all tasks. Participants made more errors with the pen than with the mouse on all tasks except the tracing task, where use of the mouse yielded a worse error rating, and the form filling task, in which no error differences were observed between devices. None of the interactions were significant. Refer to Table 2.

User preference measures

The mouse was the preferred device for the Excel dragging task, the PowerPoint manipulation task, and both cross-referencing tasks for both younger and older adults. However, participants were fairly evenly split on the PowerPoint dragging task, with 50% of younger adults and 58.33% of older adults preferring the pen to the mouse. The pen was the unanimous choice for the tracing task in both groups.

Participants were asked to rate the devices overall on a scale from 1 to 5 with 1 being the worst and 5 being the best Younger adults ($M = 4.17$) rated devices overall significantly higher than older adults ($M = 2.75$), $F(1, 20) = 14.60, p < .01, \eta^2 = .42$, and the pen received lower marks overall ($M = 2.58$) than the mouse ($M = 4.33$), $F(1, 20) = 22.27, p < .001, \eta^2 = .53$. The interaction between age and device was not significant, $F(1, 20) = 1.26, p = .27, \eta^2 = .06$. Refer to Figure 1.

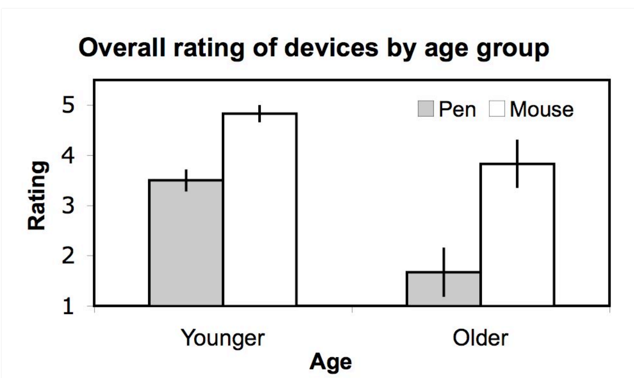

Figure 1. Overall ratings of device by age group (first device only). Error bars represent +/-1 standard error of the mean.

DISCUSSION

Performance measures

Participant expertise with devices. Consideration was given to the fact that with 10 years median experience, participants could be considered experts at using the mouse, whereas with 0 years median experience, novices at using the pen. As a result, in cases where the pen might be hypothesized to be better than the mouse, accepting the null hypothesis might reflect a positive result for the pen.

Age hypotheses. As hypothesized, older adults were generally slower than younger adults on all tasks, with the exception of the tracing task, where time was the same for both groups. Age effects for errors were less pronounced, likely reflecting a speed-accuracy tradeoff. Older adults made more errors than younger adults on the Excel dragging task, and the tracing task. Only on the first task, the Excel dragging task, did older adults make more errors as well as take more time than younger adults. As the Excel dragging task was early in the protocol it is possible that participants were still getting used to using the pen.

Device hypotheses. Taking into account the relative expertise of participants with the devices, the device hypotheses were for the most part supported by the study.

For the Excel dragging, PowerPoint manipulation, and cross-referencing document editing task, all of which require precision activity to select the correct tools or perform actions, participants were clearly more successful with the mouse. Although only on the tracing task did participants perform faster and make fewer errors with the pen, on the PowerPoint dragging task and cross-referencing form filling task, there was no difference in time performance between using the pen or the mouse. Considering that participants were novices with the pen, these results suggest that the pen supported these tasks at least as well as the mouse, and with practice, may be a better device to use than the mouse. Alternately, it may be that regardless of expertise, both devices are equally suited for the tasks. This interpretation seems likely for the cross-referencing form filling task because its properties favor both the mouse (precision work) and the pen (ballistic actions, analog to real activity). With the exception of the tracing task, however, more errors were made with the pen.

Age and device interactions. One of the goals of the present study was to investigate interactions of age and device in more realistic tasks. Previous work had found that age-related decline in performance was mitigated by use of a pen instead of a mouse, and it was hypothesized that these effects would be seen in the tasks that did not require precision actions. Such a finding would have real-world implications for potential uses of computer devices for aging populations. However, in all tasks, the null hypothesis for interactions between age and device was retained for the performance measures of time and errors. Thus the effects that were observed in elemental task studies do not appear to hold up in more complex, realistic tasks.

User preferences and qualitative indicators

While not surprising that participants unanimously preferred the pen over the mouse for the tracing task, it bodes well for the pen that a respectable portion of participants in both age groups also preferred the pen for the other two tasks that did not require highly precise actions and which had a compelling analog to a real activity (the PowerPoint dragging task and the cross-referencing form filling task).

While the pattern of device preference for tasks was generally the same for younger adults as it was for older adults, the younger adults gave higher overall ratings to the pen than the mouse, suggesting a greater willingness on the part of the younger adults to try using the pen for computing.

Limitations

Some limitations of the study included the laboratory "office" environment which may have favored the mouse, and a small participant sample resulting in inadequate power to detect subtle interactions. Also, the presence of differential carryover effects in many of the tasks suggests that for these types of studies, more practice and training is required with a novel device before it can be reliably compared to a device with which participants have experience.

Conclusions

There were three primary issues this study was intended to address: first, to compare two input devices with different interaction properties in new technology currently available to consumers; second, to include older and younger adults to see if device choice could mitigate age-related decline in performance; and third, to use goal-oriented tasks to test devices in more realistic contexts.

For the comparison of the input devices, the study supports the importance of task demands on selecting appropriate devices for productivity. Access to a greater variety of input devices would seem to be beneficial, although younger adults seemed more likely to adopt novel devices than older adults.

With regard to the factor of age, the study affirmed the robust effect of age-related decrements in performance time on computing tasks. However, benefits to older adults by use of a pen were not found in the more complex goal-oriented tasks used. Thus the value in using a specific device is likely limited to the task-related benefits.

Finally, the use of goal-oriented tasks provided a meaningful examination of the tools using contemporary software applications and more realistic contexts for the participant to evaluate the devices than they might experience in an elemental task study. While the complexity of the tasks may have obscured some of the results, future research could look at specific elements within these goal-oriented contexts, or explore a wider variety of use contexts.

REFERENCES

Adler, A., Gujar, A., Harrison, B. L., O'Hara, K., & Sellen, A. (1998). A diary study of work-related reading: Design implications for digital reading devices. *Proceedings of CHI '98: ACM Conference on Human Factors in Computing Systems*, 241-248.

Bullinger, H., Kern, P., & Braun, M. (1997). Controls. In G. Salvendy (Ed.) *Handbook of Human Factors and Ergonomics* (pp. 697-728). Mahwah, NJ: Lawrence Erlbaum.

Chaparro, A., Bohan, M., Fernandez, J., Choi, S. D., & Kattel, B. (1999). The impact of age on computer input device use: Psychophysical and physiological measures. *International Journal of Industrial Ergonomics, 24,* 503-513.

Charness, N., Holley, P., Feddon, J., & Jastrzembski, T. (2004). Light pen use and practice minimize age and hand performance differences in pointing tasks. *Human Factors, 46(3)*, 373-384.

Cohen, O., Meyer, S., and Nilsen, E. (1993) Studying the movement of high tech rodentia: Pointing and dragging. *Adjunct Proceedings of INTERCHI '93: ACM New York*, 135-136.

Greenstein, J. S. (1997). Pointing devices. In M Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction. Second, completely revised edition* (pp. 1317-1348). New York, NY: Elsevier Science.

Hinckley, K. (2003). Input technologies and techniques. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook* (pp. 151-168). Mahwah, NJ: Lawrence Erlbaum.

MacKenzie, S., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. *Proceedings of ACM CHI '91 Conference on Human Factors in Computing Systems and Graphics Interface,* 161-166. New York: ACM Press.

Meyer, A. (1995). Pen computing: a technology overview and a vision. *ACM SigCHI Bulletin, 27(3)*, 46-90.

Murata, A., & Iwase, H. (2006). Usability of touch-panel interfaces for older adults. *Human Factors, 47(4),* 767-776.

Phillips, L. M. (2003). Effect of computer input device on speed and accuracy of older adults. *Masters Abstracts International, 42* (05), 1842, (AAT 1418714).

Proctor, R. W., Vu, K. L., Pick, D. F. (2005). Aging and response selection in spatial choice tasks. *Human Factors, 47(2),* 250-269.

Rogers, W. A., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a screen or turn a knob: Choosing the best device for the job. *Human Factors, 47(2),* 271-288.

Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors, 41(3),* 388-396.

Stanton, N. A. (2003). Human error identification in human-computer interaction. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook* (pp. 371-383). Mahwah, NJ: Lawrence Erlbaum.

Vercruyssen, M. (1997). Movement control and speed of behavior. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of Human Factors and the Older Adult*. San Diego, CA: Academic Press.

Walker, N., Millians, J., & Worden, A. (1996). Mouse accelerations and performance of older computer users. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting – 1996,* 151-154.