A Comparison of Touchscreen and Mouse for Real-World and Abstract Tasks with Older Adults

by

Kailun Zhang

B.Sc., The University of British Columbia, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate and Postdoctoral Studies
(Computer Science)

The University of British Columbia
(Vancouver)

November 2015

Abstract

Touchscreens have become a mainstream input device for older adults. We compared performance of touchscreen and mouse input for older adults on both abstract and real-world pointing and dragging tasks: classic Fitts's law tasks and tasks drawn from C-TOC, a computerized cognitive test being designed for older adults. The abstract and real-world tasks were designed to require equivalent motor skills. Sixteen older adult participants completed both types of tasks using a touchscreen and a mouse. The touchscreen was faster for both task types but somewhat more error-prone. However, the speed advantage of touchscreens for abstract tasks did not translate evenly to the corresponding real-world tasks. A Keystroke-Level Model (KLM) was used to explain the different speed gains in real-world tasks by incorporating both physical and cognitive components.

As a self-administered test, C-TOC, would benefit from richer performance measures, beyond speed and accuracy, to compensate for the lack of a clinician observer who is typically present in comparable paper-based cognitive tests. We looked into the movement patterns of a real-world dragging task – the C-TOC *Pattern Construction* task – and found that older adults naturally adopted different movement patterns between devices: they tended to make shorter moves and a greater number of moves on a touchscreen than with a mouse. This indicates that careful device-based calibration will be needed for new performance metrics in computerized tests.

Preface

The study described in this thesis was conducted under the approval of the University of British Columbia (UBC) Behavioral Research Ethics Board (certificate number H09-02293 C-TOC).

Parts of this thesis appear in a conference paper manuscript

> K. Zhang, S.-H. Kim, J. McGrenere, K. Booth, C. Jacova. A Comparison of Touchscreen and Mouse for Real-World and Abstract Tasks with Older Adults.

where I am the first author. Sung-Hee Kim, Joanna McGrenere, Kellogg Booth, and Claudia Jacova helped frame and write the manuscript. Joanna McGrenere supervised the research.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANOVA Analysis of Variance

C-TOC Cognitive Test on a Computer

ID index of difficulty

KLM Keystroke-Level Model

RM-ANOVA repeated measure Analysis of Variance

Acknowledgments

I thank my supervisor Joanna McGrenere for her support, advice, guidance, valuable comments and suggestions.

I thank the C-TOC team at the UBC departments of Computer Science and Faculty of Medicine for their continuous support and feedback in this work: Claudia Jacova for her insightful suggestion regarding cognitive testing; Sung-Hee Kim for her considerate and detailed guidance throughout the work. I thank James Riggs and Charlotte Tang for their feedback.

I thank my second reader Kellogg Booth, for his valuable advice and taking the time to review and improve this work.

I thank the LUNCH and MUX group for their support: LUNCH discussants Antoine Ponsard, Kamyar Ardekani, Matt Bremer, and Mona Haraty; thesis draft reader Oliver Schneider; and pilot subjects Kamyar Ardekani, Johanna Fulda, and Hasti Seifi.

I am grateful to the funding received through the GRAND (Graphics Animation and New Media) NCE as well as the NSERC Discovery Grant program.

I thank Mike Ming-An Wu who supported me during ups and downs.

I thank my parents, Ming-Ming Zhang and Ming-Xin Zhao, for their encouragement and support.

Chapter 1

Introduction

Older adults are increasingly using computers [28], a trend that has been influenced by the commercial introduction of touchscreen devices, such as the iPad. Touchscreens have become very popular in recent years [25], in part for their ease of use and intuitiveness. They are known to require less previous experience and have been particularly welcomed by older adults [2]. Two questions arise: (1) how do touch-based devices compare with more classic devices, such as a desktop computer with a mouse, and (2) what are the relative strengths and weaknesses of these devices for the older adult population?

1.1 Extending from Abstract to Real-World Tasks

There have been many studies of mouse and touchscreen usage, but relatively few that focus on older adults. Prior studies on older adults suggest that touchscreen input is faster than mouse for pointing tasks [17], but accuracy can be noticeably worse, especially for smaller targets which require higher precision [13]. For dragging tasks, however, the literature is quite mixed: some studies found performance on touchscreens is comparable to mouse [8], but others found the mouse to be faster [30]. Further, the effect of precision level for dragging tasks has not been well studied with older adults.

Prior research comparing touchscreen and mouse, with both younger and older adults, has almost exclusively used abstract "laboratory" tasks. Fitts's law is the *de facto* standard for comparing pointing and dragging, but we wanted to know how performance on abstract Fitts's tasks translates to real-world tasks. Specifically, does the speed gain for touchscreens over mouse remain in realistic tasks that involve short movements similar to those in the abstract tasks?

The motivating context for this work is Cognitive Test on a Computer (C-TOC). C-TOC is a novel computerized test that screens for the early detection of cognitive impairment in older adults. It is currently under development. It runs in a web browser and comprises thirteen short subtests. The ultimate goal is for older adults (55+) to self-administer C-TOC using their own computing devices at home. Investigating differences in performance for older adults between a touchscreen and mouse is an important step for the C-TOC project. Making C-TOC usable with either touchscreen or mouse (it was previously mouse only) should make C-TOC more widely accessible, and it could ease test-takers' discomfort [18]. However, it is critical to identify any performance differences between the two devices so the differences can taken into account when interpreting C-TOC results – knowing accurate baseline performance is a requirement for cognition assessment.

The primary research goal had two components: (1) determine if there are differences in performance, i.e., speed or accuracy, on touchscreen vs. mouse for abstract tasks that are comparable in movement difficulty to the C-TOC tasks, and (2) understand if any performance differences that are found translate to the C-TOC real-world tasks. To achieve this we chose four C-TOC substests that have both pointing and dragging interaction for both low and high precision. We then mapped these to abstract Fitts's tasks, controlling for index of difficulty throughout. Sixteen older adult participants completed all four of the real-world C-TOC subtests as well as the abstract tasks that were deemed to be the equivalent from a motor perspective, using both touchscreen and mouse-based devices.

1.2 New Performance Metrics for C-TOC

Because C-TOC is computer based, logging test-takers' detailed interaction throughout C-TOC is easy. This type of data capture may partially compensate for the biggest disadvantage of computerized testing – the lack of observation from the human examiner who is present during standard paper-based cognitive testing. We had been curious to know what other interaction metrics, beyond speed and accuracy, might be available to evaluate participants' cognitive performance while taking C-TOC and whether these would be device sensitive.

Thus, the secondary research goal was to explore measures other than speed and accuracy that might be valuable for evaluating participants' cognitive performance while taking C-TOC, and to determine whether those measures are device sensitive. In a similar vein, we wanted to clarify any subjective differences in test-taker experience between touchscreen and mouse interfaces.

1.3 Thesis Contributions

The research contributions are as follows.

1. We replicated previous Fitts's law research for both pointing and dragging tasks, reinforcing its applicability to older adults: touchscreen is faster than mouse, but less accurate in high-precision tasks.

2. We are the first, to our knowledge, to systematically extend the comparison between touchscreen and mouse beyond abstract tasks to a real-world context: speed and accuracy differences between devices don't translate evenly from abstract to real-world tasks due to the cognitive component involved in C-TOC. We analyzed the speed gain difference between abstract and C-TOC tasks using the Keystroke-Level Model (KLM), which supported the data gathered.

3. We uncovered considerably different movement patterns between devices in a real-world dragging task: touchscreen yields nearly 50% more moves compared to the mouse, but this did not translate into differences in total task completion time between the two devices. We further investigated the difference in movement patterns by coding participants' individual dragging moves into a set of categories, and found that participants, instead of making just single movements, often separate a move into multiple shorter moves on a touchscreen.

4. We show a relationship between age, manual dexterity and performance,

which may explain older adults' strong preference for the touchscreen in pointing tasks and lack of such preference in dragging tasks.

1.4 Overview of the Thesis

Previous work relevant to the research is summarized in Chapter 2. Chapter 3 discusses design considerations in selecting abstract and real-world C-TOC tasks for an experiment we conducted to investigate the research questions. Chapter 4 describes the experimental methodology, followed by a presentation of the results of the experiment in Chapter 5. Chapter 6 interprets the results and offers a KLM-style analysis of the findings for speed of task completion. Chapter 7 summarizes the findings in the thesis and discusses directions for future work.

Chapter 2

Related Work

2.1 Older Adults and Device Comparisons

There has been research done about how to better support older adults' computer usage that evaluated novel or less common input devices, including light pen [4], eye-gaze [16, 22], EZ ball [30], and rotary encoder [20]. Given the C-TOC context, we take a pragmatic approach and compare two mainstream input devices: mouse and touchscreen.

It is well known that the relative advantage of an input device depends on task and context [3, 4, 20]. Different tasks or even different contexts may require different types of interaction. Findlater et al. [8] found that older adults were 35% faster using touchscreen compared to mouse, but speed gain was much bigger in some interaction types (pointing and crossing) than others (dragging and steering). We focus on pointing and dragging, the only two interaction techniques used in C-TOC.

Performing pointing tasks on a touchscreen is known to be significantly faster than using a mouse but much more error-prone, not only for the general population [5, 9, 21] but also for older adults: Ng et al. [17] found that pointing on a touchscreen was 100% faster than with a mouse for older adults. However, other research has shown that accuracy suffers as a result. Touchscreens are especially inaccurate for small target sizes. Kobayashi et al. [13] found a target width of 30px was too small for older adults to point to with a finger. The error rate for a target of this size was 13.6% for iPad and 39% for iPod. Performance did not improve even after a week of practice. However, the same high error rate was not found for target sizes just a bit larger, indicating that performance may not degrade smoothly.

For dragging tasks, studies have had inconsistent results comparing touchscreen and mouse. Findlater et al. [8] found comparable dragging times for older adults, but Wood et al. [30] found touchscreen was 40% slower, also for older adults. However, Wood et al.'s 32px icon size is too small according to Kobayashi et al.'s standard [13].

All studies comparing computer input devices, with older adults as the participants, have exclusively used abstract tasks that have little or no cognitive component, with one exception. Rogers et al. [20] used an Entertainment System Simulator to evaluate performance of a touchscreen and a little-known device, a rotary encoder. Neither device was a clear winner.

2.2 Effect of Age and Dexterity on Input Device Performance

Aging typically affects performance with input devices negatively, due to the various functional declines associated with aging, although the degree at which aging affects performance differs across devices. For pointing tasks, aging has a lesser effect on the task performance using a touchscreen than using a mouse [11, 17]. To our knowledge, no previous research has studied the effect of aging for dragging tasks.

One of the functional abilities closely related to aging is manual dexterity. Previous studies have tried to isolate the effect of manual dexterity on input device performance. Jin et al. [12] reported that lower manual dexterity led older adults to spend significantly longer time performing pointing on a touchscreen. The effect of manual dexterity, however, has not been studied in the context of dragging tasks, nor compared between different input devices.

2.3 Computerized Cognitive Tests

Standard practice for cognitive assessments is paper-based testing in clinical settings [14]. Attempts have been made to develop computer-based cognitive tests for older adults [7, 19, 27]. All these cognitive tests are administered only on a specific platform with only one type of input device. There is no research comparing test performance across devices.

None of the computerized cognitive tests are self-administered or taken at home [29, 31]. Most computer-based tests are adaptations of the paper-and-pencil versions of neuropsychological tests [23], where observations from human examiners complement the test scores [14, 26]. Because C-TOC is being designed to be self-administered, it will not have the benefit of a human examiner. Finding ways to make up for the missing observational data and complement the scores is a challenge.

Chapter 3

Task Design

We investigated two types of tasks: abstract tasks and real-world C-TOC tasks, each spanning two interaction types (pointing and dragging) and requiring both high and low precision. C-TOC tasks were drawn from actual C-TOC subtests. Abstract tasks were traditional Fitts's law tasks that were chosen to approximately match the precision required by the corresponding C-TOC tasks. We used the idea of task precision from Fitts's law, namely index of difficulty (ID), to estimate task precision of the selected C-TOC tasks. ID is calculated from target width (W) and movement amplitude (A). We used the Shannon formulation, $ID = \log_2(A/W + 1)$, recommended by MacKenzie [15]. We start by explaining C-TOC tasks and then how we estimated the ID for each C-TOC task to determine the task precision levels for the abstract tasks.

3.1 Real-World C-TOC Tasks

We selected four C-TOC subtests for the experiment: *Picture-Word Pairs*, *Arithmetic*, *Sentence Comprehension*, and *Pattern Construction*. Figure 3.1 shows how each subtest corresponds to a precision level (low or high) and task type (pointing or dragging). C-TOC scoring for subtests depends on either accuracy alone, or a combination of accuracy and speed. In the performance analysis, we report accuracy and speed individually, instead of reporting a C-TOC score. The subtests and estimated task precisions are described in the following subsections.

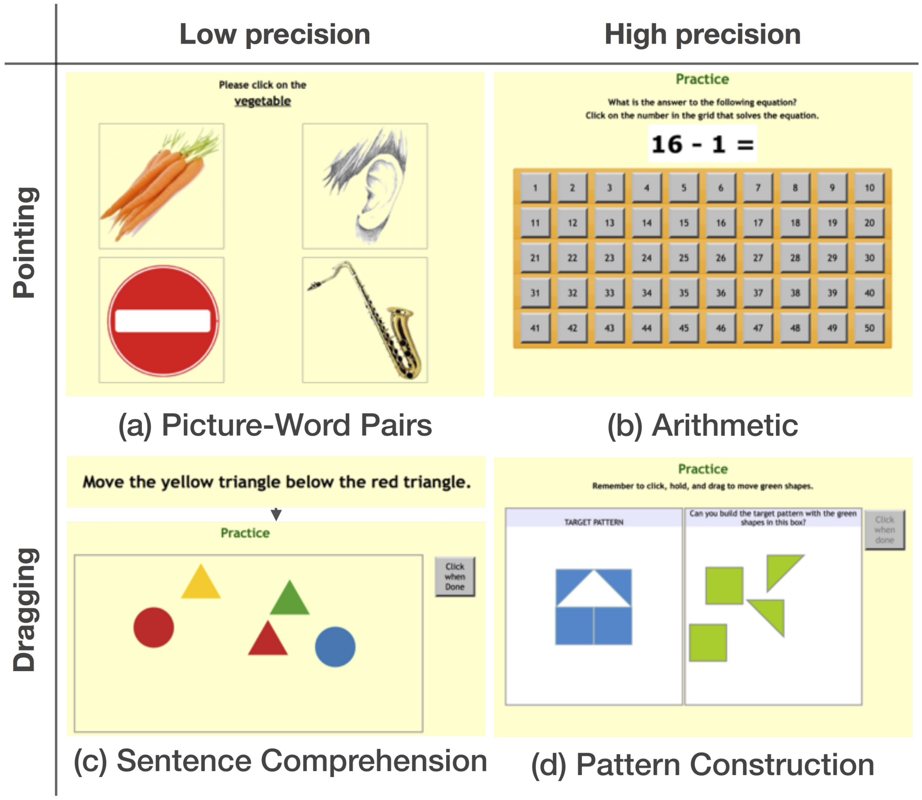

Figure 3.1: The four C-TOC subtests used in the experiment. Each subtest corresponds to a task type (pointing or dragging) and a precision level (low or high).

3.1.1 Low-Precision Pointing: Picture-Word Pairs

Picture-Word Pairs (Figure 3.1a) is a memory encoding task. The participant is presented with four images and an instruction such as “Please click on the vegetable.” The participant must click/tap on one of the four images, which ends the trial. Each trial starts with the participant clicking an “OK” button in a pop-up window in the middle of the screen. This ensures that the mouse cursor or the finger always starts from the same position. Task precision is $ID \approx 1.0$, which was calculated by the width of the images (250px) and amplitude, which is the distance between the start and the end point in the task (250px).

3.1.2 High-Precision Pointing: Arithmetic

Arithmetic (Figure 3.1b) tests numeracy with simple arithmetic problems and four basic operators ($+ - \times \div$). To answer, a grid of clickable buttons corresponding to numbers from 1 to 50 is provided. Each trial starts with the participant clicking an "OK" button in a pop-up window in the middle of the screen (the same as for the *Picture-Word Pairs* subtest), and ends when the participant clicks one of the buttons. Task precision is $ID \approx 2.5$, based on the width of the number buttons (70px) and the distance between the start and end point for the task (250px-400px).

3.1.3 Low-Precision Dragging: Sentence Comprehension

Sentence Comprehension (Figure 3.1c) tests short-term memory. It has two stages: (1) memorize the instruction given on the screen, such as "Move the yellow triangle below the red triangle" and then click the "Next" button, which will transit to the second stage and move on to a new screen; (2) then, among the movable shapes, drag the shapes as per the instruction and then click the button "Click when Done." Two types of time periods were measured: (1) task completion time, which is the period between clicking "Next" and clicking the "Click when Done" button, and (2) times for each dragging movement. Task precision is calculated based on width of the intended target zone (Figure 3.2a) and movement amplitude. The width of the movable shapes is 80-100px, the width of intended target zones is 200-400px, and movement amplitude is 200-400px, varying across trials. The width of intended target zones and the amplitudes were verified in pilot tests. Task precision is $ID \approx 1.0$. Older adults have large variance in performance [2], so the perceived width of intended target zone can vary, resulting in discrepancies estimating task precision.

3.1.4 High-Precision Dragging: Pattern Construction

Pattern Construction (Figure 3.1d) is a visuospatial test. The participant is asked to drag a set of movable shapes to match a reference target pattern that remains visible throughout the test. Shapes can be translated but not rotated. Two types of time periods were measured: (1) task completion time that starts when the screen appears showing the target pattern and movable shapes, and ends when the participant

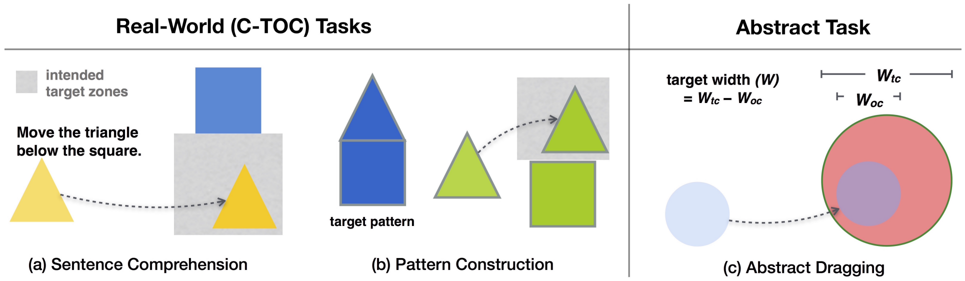

Figure 3.2: For real-world tasks, we defined an intended target zone. The intended target zone size is larger in the low-precision dragging task (a) *Sentence Comprehension* compared to the high-precision dragging task (b) *Pattern Construction*. The abstract dragging task (c) is adjusted to be comparable to the real-world tasks, in which participants were asked to drag the blue object circle (W_{oc}) fully in into the red target circle (W_{tc}).

clicks the "Click when Done" button, and (2) times for each dragging movement. Due to flexibility in constructing patterns with multiple objects, the precision required for a specific dragging movement could be low or high, but the maximum precision is estimated as $ID \approx 2.5$. The width of movable shapes is 80-160px. For high-precision movements, the intended target zone is 0-30px wider than the movable shape. Movement amplitude is up to 150-200px.

3.2 Abstract Tasks

Abstract tasks were multi-directional pointing and dragging tasks, implemented based on ISO:9241-400 [10] (see Figure 3.3). For pointing, the participant is asked to click or tap on a target object.

For a dragging task, we modified the standard Fitts's law dragging task. The participant is asked to drag an object circle fully into, as opposed to partially overlapping with, a target circle to successfully complete the task (Figure 3.2c). The circumference of the target circle highlights in green once the object circle is fully within the target circle. The modification was to better mimic C-TOC dragging tasks in which participants drag an object shape into an intended target zone. For dragging, the target width (W) is defined as the difference between the object circle width (W_{oc}) and target circle width (W_{tc}).

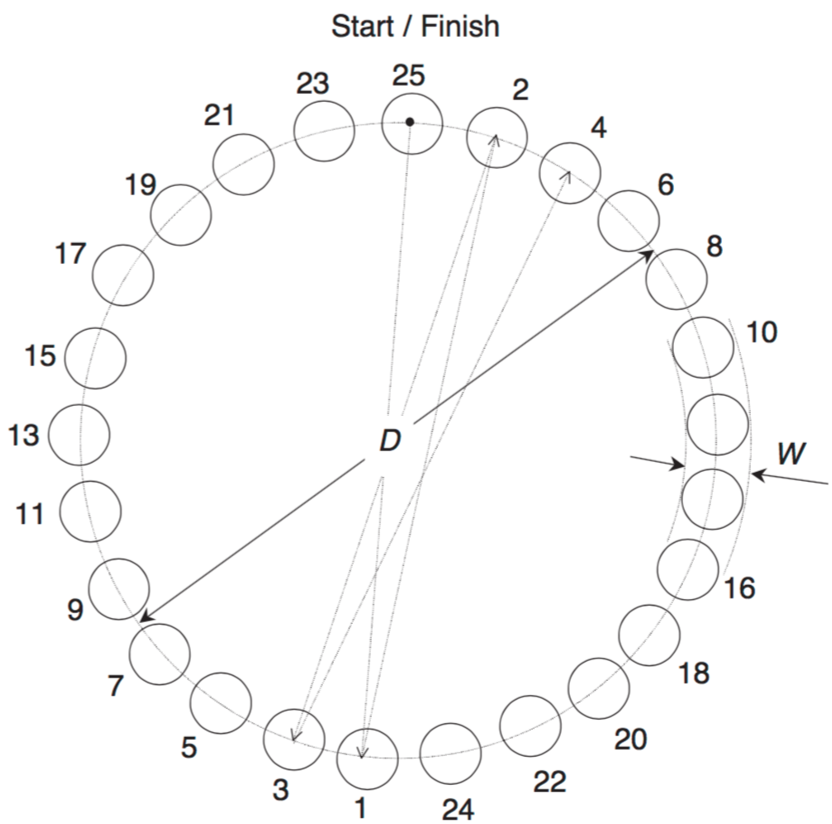

Figure 3.3: Paradigm for multi-directional pointing and dragging tasks. Figure copied from Soukoreff and MacKenzie [24].

For both the pointing and dragging abstract tasks, amplitude (A) is 250px so the largest target circle (approximately 250px) fits within the iPad screen (768px on the short edge). We determined W values from the amplitude 250px and the set of IDs that reflect task precisions in C-TOC subtests. We had two object widths (50px and 80px) in the dragging tasks so we could test for an effect of object width.

We chose three task precisions: IDs of 1.0, 2.5, and 3.0. The first two approximated the precisions in the low- and high-precision C-TOC tasks. $ID = 3.0$ was included to cover a wider range for trend analysis. An ID higher than 3.0 was excluded because target width would be too small for touchscreens [13]. Although all *ID*s are considered low precision compared to typical *ID* values of 2 to 8 for abstract tasks [10], interfaces designed for older adults typically require lower task precision (i.e., larger targets) compared to interfaces for the general population. For each precision, we added two variants, 0.9 and 1.1 for $ID = 1.0$, etc., for a total of 9 *ID*s (*A*-*W* pairs), to allow flexibility in estimating task precision and to ensure sufficient power for regression modeling.

Chapter 4

Methods

This chapter discusses the detailed methodology used in the experiment. Tasks have already been described in Chapter 3 Task Design.

4.1 Design

The experiment included four factors: task, input device, interaction type, and task precision. Each participant completed four abstract task conditions: 2 (touchscreen vs. mouse) $\times$ 2 (pointing vs. dragging), as well as eight C-TOC subtest conditions: 2 (touchscreen vs. mouse) $\times$ 2 (pointing vs. dragging) $\times$ 2 (precision levels). Each abstract task condition contained nine precision levels (0.90, 1.00, 1.10, 2.40, 2.50, 2.60, 2.90, 3.00, 3.10) that were fully randomized across six repetitions in pointing and three repetitions in dragging with two object sizes[1]. Optional break times were evenly distributed within each condition. The orders of task, input device, and interaction type were fully counterbalanced. For C-TOC tasks, the order of task precision was fully counterbalanced and we fully randomized two isomorphic sets of trials to ensure that participants did not see the same trials in both the touchscreen and the mouse conditions.

Although the experimental design included four factors, our primary interest was to understand the effect of device on the factors of task (abstract vs. real-

[1]Dragging tasks have two object sizes with each target width (W) whereas pointing tasks have only one object size in each W. To achieve the same total number of trials per task, dragging tasks have half the number of repetitions compared to pointing tasks.

world), interaction type (pointing vs. dragging), and precision level. We were not interested in directly comparing interaction types to each other (it is well known that pointing is faster than dragging) nor in directly comparing the C-TOC tasks to each other (they are very different).

4.2 Procedure

After signing a consent form (Appendix A.2), a participant completed a demographic questionnaire about age, gender, motor and visual impairments, and frequency of computer usage (Appendix A.3.1). The frequencies of touchscreen and mouse usage were collected using 5-point Likert scales as part of the questionnaire, followed by administration of the Purdue Pegboard Test and the Snellen Vision Test to measure manual dexterity and eyesight.

Participants alternated between abstract and C-TOC tasks and the order was counterbalanced. The order of precision level for abstract tasks was randomized; the precision level for C-TOC subtests was determined by the task and was counterbalanced. Participants used one device for all tasks before switching to the other device. Within each device, they first performed all tasks of a single interaction type (pointing or dragging) before tasks of the other interaction type. They had practice trials throughout and were offered breaks between each task.

After completing all trials, the Purdue Pegboard test was administered a second time to check for fatigue. A session concluded with an interview asking for the preferred device for each task and why it was preferred (Appendix A.3.2). Total duration of a study session was approximately 1.5 hours.

4.3 Participants

Sixteen people (10 female) ages $57-88$ years ($M = 71.81, SD = 9.60$) participated in the study, all right-handed, none with any diagnosed cognitive impairment. We used the participants' score on the first Purdue Pegboard Test as an indication of their manual dexterity (see Table 4.1 for a summary of detailed scores).

Fourteen participants reported no conditions that would affect motor ability. Two reported having arthritis, but their Pegboard results were better than the predicted scores for senior adults of their age [6], so we included their data.

Table 4.1: Summary of Purdue Pegboard score for all participants

	Min	Max	Mean	Std. Dev.
Right hand	8	17	12.25	2.46
Left hand	8	17	11.38	2.47
Both hands	5	15	9.63	2.55
Assembly	16	40	25.44	7.31

No participant had a significant drop in Pegboard Test score after completing the experiment, indicating fatigue was not an issue. Results from the eyesight test showed no visual deficiency for any participant that might affect performance. All participants but one owned a desktop or a laptop with a mouse. Half (eight out of sixteen) of the participants had access to a touchscreen device.

4.4 Apparatus

The experiment was implemented in JavaScript, HTML, and PHP and built with the Raphaël vector graphics library. It ran on an iPad 4th-generation (touchscreen condition) and a 13-inch MacBook Pro with a Logitech Wireless Mouse M310 (mouse condition). Both devices had retina displays with resolution 1024×768 pixels (iPad) and 1280×800 pixels (MacBook Pro). The experiment was run in the Safari browser on both devices (version 8.0 on iPad under iOS 8.3, 8.0.4 on MacBook under OS X Yosemite). The iPad was set in landscape orientation and tilted at a fixed 20-degree angle.

During the experiment, we recorded the screen of the devices, participants' hands interacting with the touchscreen, and audio of the interview sessions.

4.5 Hypotheses

We use speed and error as measures for evaluating performance. For abstract tasks, hypotheses are derived from previous findings for abstract tasks. For real-world C-TOC tasks, we hypothesized that time differences between input devices would be washed out by the cognitive component involved in the test (**H2-a**). We also thought participants would try for best performance in accuracy in a cognitive test, thus there would be no difference in accuracy between devices (**H2-b**).

H1. Replication of previous work on abstract tasks:

(a) Pointing is faster on touchscreen.

(b) For a dragging task, there is no difference in speed between devices.

(c) Error rate is higher in tasks with higher *ID*s in both pointing and dragging tasks.

H2. In real-world tasks, between touchscreen and mouse:

(a) There is no difference in speed.

(b) There is no difference in accuracy.

H3. Participants' subjective experience:

(a) Preference for touchscreen will be stronger in pointing tasks than in dragging tasks.

(b) Participants will prefer touchscreen over mouse for both abstract and C-TOC tasks.

Beyond performance and preference, we were interested to explore if there were other qualitative differences in experience between touchscreen and mouse for older adults.

Chapter 5

Results

We begin by comparing results for the abstract tasks and C-TOC tasks, followed by analyzing behavioral difference observed in the C-TOC *Pattern Construction* task, and then the subjective preferences for each device. Lastly, we present a post-hoc analysis investigating the influence of dexterity and age on speed for both devices.

Throughout the analyses, we determined generalized eta-square (η_G^2) for effect size using Bakeman's [1] suggestion, with the interpretation of .02 as small, .13 as medium and .26 as large effect size. Post-hoc comparisons are adjusted using the Bonferroni correction. Given the disparity in error rates and the low $A:W$ ratio, following MacKenzie [15], we determined effective target width (W_e) and effective movement amplitude (A_e) for each A-W pair.

5.1 Performance of Abstract Pointing and Dragging

We describe the outlier removal process, and then the results for accuracy, speed, and the regression analysis.

5.1.1 Outlier Removal

Two criteria were used for detecting spatial outliers. We eliminated trials in which movement was less than half the trial amplitude and we eliminated trials with movement greater than three standard deviations from the mean, where means and standard deviations were for each subject, *device*, *task*, and *target width*. Outliers accounted for 1.9% of all trials. Speed and error analyses in the following sections exclude all outliers.

5.1.2 Pointing Task

We used a 2×9 repeated measure Analysis of Variance (RM-ANOVA): *input device* by *task precision* (9 *ID*s).

Speed

The decrease in speed as task precision increased was significantly larger for the mouse than for touchscreen: a *device* $\times$ *task precision* interaction dominated ($F_{8,120} = 31.16, p < .001, \eta_G^2 = .33$). Overall, touch was faster: main effect of *device* ($F_{1,15} = 75.90, p < .001, \eta_G^2 = .73$). As *task precision* increased, speed decreased (Figure 5.1, left): main effect of *task precision* ($F_{8,120} = 72.25, p < .001, \eta_G^2 = .53$). Post-hoc tests revealed touchscreen had no significant increase on pointing time across all *ID* levels, and it was always faster than mouse at the same *ID* level ($p < .001$). Pointing using a mouse for *ID* higher than 2.4 was slower than pointing using touchscreens regardless of *ID* ($p < .001$).

Accuracy

Overall error rate was 5.59% for the pointing task, but it was dependent on *input device* and *task precision*: a *device* $\times$ *task precision* interaction dominated ($F_{8,120} = 7.56, p < .001, \eta_G^2 = .18$). There was a relatively constant error rate across precision levels for the mouse, but an abrupt increase in error rate as *task precision* increased for touchscreen. Touchscreen was particularly inaccurate for high-precision tasks in which target width was around 4mm (33-40px). Overall, touch had more errors: main effect of *device* ($F_{1,15} = 20.28, p < .001, \eta_G^2 = .20$). As precision increased, so did the rate of errors (Figure 5.2, left): main effect of *task precision* ($F_{8,120} = 12.50, p < .001, \eta_G^2 = .23$). Post-hoc tests showed the error rate for touchscreen pointing in high-precision ($ID \geq 2.9$) was always significantly higher than the error rate in (1) any other precision level of touchscreen pointing and (2) all mouse pointing regardless of task precision ($p < .001$).

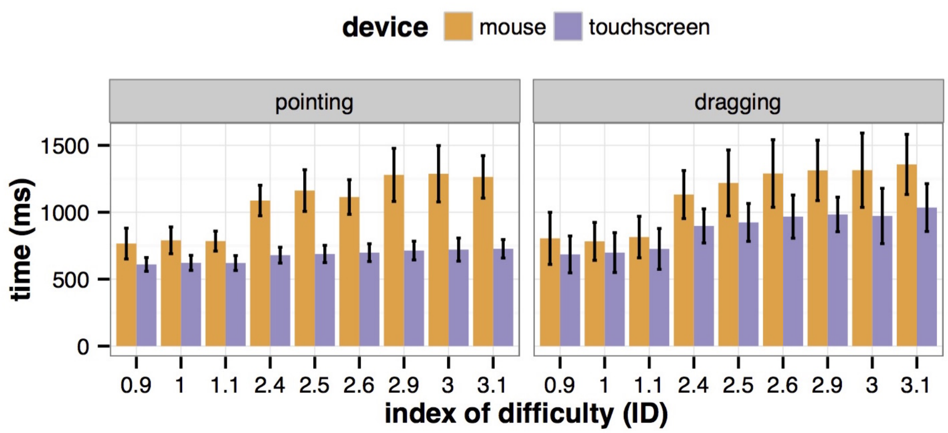

Figure 5.1: Speed for abstract tasks by *device* and *task precision*. Error bars show the 95% confidence interval.

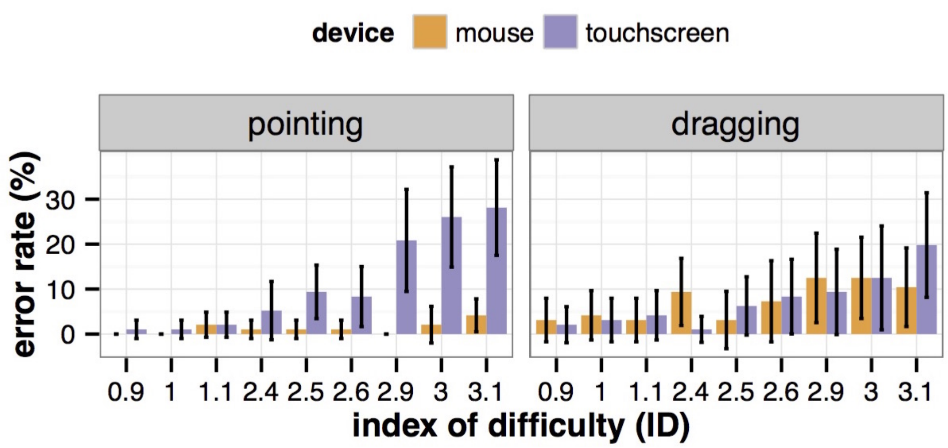

Figure 5.2: Error rate for abstract tasks by *device* and *task precision*. Error bars show the 95% confidence interval.

Table 5.1: Regression analysis for each device-task combinations. Throughput is calculated by $1/b$, where b is the slope of the model.

Device	R^2	Regression Coefficients Intercept (ms)	Slope (ms/bit)	Throughput (bits/s)
Pointing				
Mouse	0.97	488	198	5.1
Touchscreen	0.92	412	93	10.8
Dragging				
Mouse	0.97	472	234	4.3
Touchscreen	0.86	414	159	6.3

Regression Analysis

A regression was performed of time on the effective index of difficulty (ID_e) that had been re-computed from W_e and A_e. As expected, touchscreen was much more efficient (107%) than the mouse in pointing tasks in *ID* ranges from 1-3, with throughput of touchscreen (10.8 bits/s) double the throughput of mouse (5.1 bit/s). The R^2 values are reported in Table 5.1.

5.1.3 Dragging Task

We used a $2 \times 2 \times 9$ RM-ANOVA for factors *input device*, *object width*, and *task precision*.

Speed

Similar to pointing, the decrease in speed as task precision increased was greater for mouse than for touchscreen: a *device* $\times$ *task precision* interaction dominated ($F_{8,120} = 5.60, p < .001, \eta_G^2 = .07$). Overall, touch was fastest: main effect of *device* ($F_{1,15} = 18.84, p < .001, \eta_G^2 = .27$). As *task precision* increased, speed decreased: main effect of *task precision* ($F_{8,120} = 61.18, p < .001, \eta_G^2 = .45$). (See Figure 5.1, right.) Larger objects yielded faster speed (975ms) compared to smaller objects (1016ms): main effect of *object width* ($F_{1,15} = 5.35, p = .035, \eta_G^2 = .01$). There was no significant interaction between object width and the other two factors. Post-hoc tests revealed touchscreen had no significant increase of time once $ID \geq 2.4$. Touchscreen and mouse had comparable speed for $ID \leq 1.1$, but significantly faster speed using touchscreen after $ID \geq 2.4$ ($p < .001$). Dragging using a mouse for *ID*s higher than 2.5 is slower than dragging using touchscreen regardless of the *ID* ($p < .001$).

Accuracy

Overall error rate was 4.24% for dragging tasks. Error rate increased as *task precision* increased, independent of input device: a main effect of *task precision* ($F_{8,120} = 4.89, p < .001, \eta_G^2 = .08$). (See Figure 5.2, right.) There was also an interaction between *object width* and *device* ($F_{1,15} = 6.37, p = .02, \eta_G^2 < .01$) and an interaction between *task precision* and *device* ($F_{8,120} = 2.41, p = .02, \eta_G^2 = .02$). However, both interactions had very small effect sizes, thus need careful interpretation. Unlike the pointing task, we did not find a dramatic increase in error rate for high-precision tasks on the touchscreen, possibly because during each dragging move, participants could continuously see and adjust the circle they grabbed with their finger until it reached the desired position, whereas each pointing move was a task with a single and instant attempt.

Regression Analysis

Performance was 46% more efficient for touchscreen than with mouse in dragging tasks with *ID* ranges from 1-3. Throughputs for mouse and for touchscreen were 4.3 bits/s and 6.3 bits/s, respectively. Similar to pointing tasks, we computed ID_e according to the adjusted W_e and A_e. There were high correlations between time and ID_e for mouse ($R^2 = 0.97$). R^2 for touchscreen was slightly lower at 0.86 (see Table 5.1).

5.1.4 Summary

Pointing was significantly faster on a touchscreen compared to a mouse (**H1-a *supported***). Dragging had less difference between devices than did pointing, but contrary to previous studies, dragging on touchscreen was significantly faster than mouse (**H1-b *not supported***). Error rate was higher with increasing precision for both tasks, with the "fat finger" problem apparently affecting pointing when using touchscreen for high precision tasks (**H1-c *supported***).

5.2 Performance of C-TOC Tasks

Speed and accuracy results for C-TOC tasks were examined using a one-way RM-ANOVA with input device as the within-subject factor. Each C-TOC task was analysed independently.

5.2.1 Speed

For low-precision pointing (*Picture-Word Pairs*), participants performed 32% faster on touchscreen than with mouse (1696ms vs. 2246ms, $F_{1,15} = 9.9, p = .006, \eta_G^2 = .40$). For high-precision pointing (*Arithmetic*), there was a trend with large effect size that using touchscreen was about 12% faster than using a mouse (5602ms vs. 6273ms, $F_{1,15} = 3.8, p = .069, \eta_G^2 = .20$).

For both time measures in the dragging tasks, mean time was always lower on touchscreen compared to mouse, but not all comparisons were significant. For low-precision dragging (*Sentence Comprehension*), task completion time was significantly faster on touchscreen compared to mouse (9.7s vs. 10.9s, $F_{1,15} = 4.76, p = .04, \eta_G^2 = .24$). But there was no significant effect of device in duration of individual dragging moves (1.2s vs. 1.4s, $F_{1,15} = 1.98, p = .18, \eta_G^2 = .12$). For high-precision dragging (*Pattern Construction*), it was the opposite. Duration of individual dragging moves was significantly faster on touchscreen (1.2s vs. 2.2s, $F_{1,15} = 18.80, p = .001, \eta_G^2 = .63$), with a 83% increase of time on mouse. As will be explained in Section 5.3, we found that participants made moves of shorter distance on touchscreen and longer distance with the mouse. When we take into account distance moved, the actual speed difference between devices is reduced to 53%. Despite this difference in speed, there was no significant difference in total task completion time (both 62s, $F_{1,15} < 0.01, p = .99, \eta_G^2 < .01$).

In Chapter 6, we provide a possible explanation based on a KLM-style analysis of why some time measures were significant, but others were not.

5.2.2 Accuracy

High-precision pointing (*Arithmetic*) was the only C-TOC task that had a significant difference of accuracy between devices (90% for mouse vs. 82.5% for touch, $F_{1,15} = 5.87, p = .028, \eta_G^2 = .28$). There were no significant differences for *Picture-Word Pairs* ($F_{1,15} = 1, p = .33, \eta_G^2 = .06$), *Sentence Comprehension* ($F_{1,15} = 0.03, p = .86, \eta_G^2 < .01$) or *Pattern Construction* ($F_{1,15} = .62, p = .44, \eta_G^2 = .04$).

5.2.3 Summary

H2 was not supported: participants performed faster on all four C-TOC tasks using touchscreen compared to using mouse, although the touchscreen speed gain varied in magnitude compared to that in the abstract tasks (**H2-a *not supported***). One of the four subtests (*Arithmetic*) had a lower error rate on touchscreen (**H2-b *partially supported***).

5.3 Movement Patterns in C-TOC Dragging Tasks

Beyond time and error, computerized cognitive tests (unlike their paper counterparts) have the possibility to infer test-takers' cognitive ability from the rich interaction log data. However, there is not much known about what potential measures might be indicative of cognitive ability. We make a first attempt at this for the C-TOC dragging task *Pattern Construction*. Due to the considerable flexibility in completing the task, *Pattern Construction* exhibited multiple behaviors because participants could use different strategies to construct a target pattern.

We observed that participants seemed to make more dragging moves with the touchscreen than with a mouse. Analysis of the data showed that there was indeed a significant difference across devices (mean of 24 vs. 15, $F_{1,15} = 13.54, p = 0.002, \eta_G^2 = .47$).

Participants also had shorter moving distance with the touchscreen compared to the mouse (50px vs. 106px, $F_{1,15} = 19.19, p < .001, \eta_G^2 = .56$).

We wondered why participants performed more dragging moves on a touchscreen than with a mouse. We examined the log data and video recording for *Pattern Construction* and generated a set of categories for the dragging moves. We coded individual dragging moves into these categories to see if the classification would reveal any differences in movement pattern between devices.

5.3.1 Coding Method

Each logged dragging move was matched with its paired screen and hand movement recordings. Based on the video, we added a *fail-to-grab* category, which was an important movement that was not always captured by the log data.

The author coded six trials for multiple participants on both devices and designed the coding scheme. A second rater used the scheme and independently coded the same six trials. The inter-rater reliability was found to be good, with Kappa = 0.83. The two raters slightly modified the coding scheme after validation. The author then coded the rest of the trials. We selected two trials from *Pattern Construction*, one more complex than the other. In total, 64 trials were coded (2 trials $\times$ 2 devices $\times$ 16 participants).

5.3.2 Classification of Types of Moves

We classified each dragging move into one of the following nine categories:

1. *Target-oriented Move* is when participants move a shape to a specific target position. There are three types of moves under this category: *single move, sub-move* and *precision adjustment*.

   (a) *Single Move* is when participants move a shape directly to the target position.

   (b) *Sub-move* is a step in a sequence of two or more steps that together move a shape to the target position.

   (c) *Precision Adjustment* is a move for fine-tuning the precise location of a shape that was largely already in target position (and is not in a sequence of *sub-move*).

2. *Trial & Error Move* is when participants attempt to move a shape to a target position, realize that the position is incorrect before releasing, and either attempt at a new target position or move the shape aside.

3. *De-construction Move* is when participants move one shape out of the already built pattern.

4. *Make-way Move* is when participants move a shape away from its current position to make room for other shapes.

5. *Rotation Attempt* only happened on touchscreen. It is an action in which participants drag the mouse or their finger in a circular trajectory with the intention to rotate a shape. This is often accompanied with verbal articulation.

6. *Accidental Click* is when participants click unintentionally.

7. *Constrained Move* is when participants try to move a shape beyond the canvas boundary.

8. *Unknown Move* is a move logged by the system whose intention could not be inferred.

9. *Fail-to-grab* is when participants attempt to grab a shape with mouse cursor or finger, but fail to do so.

Table 5.2 gives a summary of the total number of moves by category across devices. Note that *accidental click* on touchscreens was under-reported because it would not be logged because it would not change the state of the interface.

Among all the computer-logged moves, *sub-moves* contributed the most towards the high count of dragging moves on the touchscreen. Participants were more likely to separate a single target-oriented move into smaller consecutive moves (*sub-moves*) on a touchscreen device compared to mouse, resulting in a shorter distance on touchscreen and longer distance with a mouse. They also tended to make more *single moves* and *trial & error* using mouse.

We found significantly more *fail-to-grab* moves on touchscreen. Though there were consistent *fail-to-grab* moves observed across all participants, some participants had very different attitudes towards this type of move. Participants not used to the touchscreen got especially annoyed and anxious if they could not grab a shape. Yet, participants with touchscreen experience reported they did not mind it at all.

Table 5.2: Summary of the total number of each type of move for each of the devices.

Type of Moves		Total Number of Moves			
		Mouse	Touch	ANOVA	
Single Move		273	205	$p = .09$	$\eta_G^2 = .17$
Sub-move		45	376	$p < .001$	$\eta_G^2 = .63$
Precision Adjustment		93	108	$p = .63$	$\eta_G^2 = .02$
Trial & Error		38	9	$p = .06$	$\eta_G^2 = .21$
De-construction		14	7	$p = .22$	$\eta_G^2 = .10$
Make-way Move		53	70	$p = .42$	$\eta_G^2 = .04$
Rotation Attempt		0	4	$p = .16$	$\eta_G^2 = .13$
Accidental Click		18	1	$p = .06$	$\eta_G^2 = .21$
Constrained Move		19	1	$p = .007$	$\eta_G^2 = .39$
Unknown		19	31	$p = .26$	$\eta_G^2 = .08$
Fail-to-grab	Logged	6	24	$p < .001$	$\eta_G^2 = .56$
	Unlogged	15	109		
Total Number of Logged Moves		565	836	$p = .002$	$\eta_G^2 = .47$

The move classification for dragging revealed some interesting issues about device affordances. Some participants performed the rotation gesture (*rotation attempt*), but only for those who received the touchscreen conditions before the mouse conditions, suggesting touchscreen is a natural device to afford complex gestures, such as rotation. Participants also performed significantly more *constrained moves* with a mouse, trying to move shapes out of the canvas boundary.

To summarize, we found participants had considerably different movement patterns across the two devices in *Pattern Construction*. Of particular note is that on a touchscreen, participants tended to make *sub-moves*, resulting in a higher number of moves but shorter distances in each move. The reverse was observed when participants used a mouse.

5.4 Subjective Preference

A summary of participants' preference of device by interaction type (pointing and dragging) and task type (abstract and real-world) is presented in Table 5.3.

Table 5.3: Participants' subjective preference of device by task. $N = 96$.

	Tasks	Mouse	Touch	Tie
Pointing	Abstract	0	14	2
	Real-World Low Precision	4	10	2
	Real-World High Precision	1	13	2
Dragging	Abstract	3	10	3
	Real-World Low Precision	2	5	9
	Real-World High Precision	7	6	3
	Totals	17	58	21

For the analysis, we excluded counts for a tie (no preference of device). We first looked into whether participants had different device preferences for pointing and dragging tasks by collapsing votes across task type. The Chi-Square test revealed preference did differ by interaction type, $\chi^2_{(1,N=75)} = 4.50, p = 0.03$. Participants expressed a much stronger preference for touchscreen over mouse for pointing tasks (37 vs. 5, with 6 ties) compared to dragging tasks (21 vs. 12, with 15 ties).

We were also interested to know if participants had different device preferences for abstract versus real-world tasks. We similarly collapsed the votes across interaction type. A Chi-Square test revealed no difference in preference of device for each type of task, $\chi^2_{(1,N=75)} = 2.01, p = 0.15$[1]. The collapsed votes for both tasks indicate that participants expressed a strong preference for touchscreen over mouse (real-world: 34 vs. 14, with 16 ties; abstract: 24 vs. 3, with 5 ties), thus there was no difference in device preference based on task.

In self-reports, participants preferred touchscreen because it was "fast", "direct", "intuitive to use", and "easier to point". In tasks where participants preferred mouse over touchscreen, the main reasons were the "high precision" of the mouse

[1]The analysis for pointing vs. dragging is significant, but abstract vs. real-world is not. In fact, the ratios are similar between the two (37 vs. 5 approx. 7 : 1; 21 vs 12 approx. 2 : 1) for the first analysis, and (34 vs. 14 approx. 2.5 : 1 ; 24 vs. 3 approx. 8 : 1) for the second. Yet the first is significant and the second is not. We followed-up by running the Fisher's exact test which is another non-parametric distribution test (similar to Chi-square). The outcome for the abstract vs. real-world comparison was $p = .09$, thus borderline significant. Altogether this points to the need for further research on subjective preferences for devices for different types of tasks

cursor, "no occlusion of finger on screen", and familiarity with a mouse. A tie in the preference for device occurred when participants reported that the cognitive workload of the task was high, as in *Sentence Comprehension*.

For the third hypothesis, preference towards touchscreen is stronger in pointing tasks than in dragging tasks (**H3-a *supported***). For both abstract and real-world tasks, there was no difference in terms of preference: the vote for touchscreen was higher (**H3-b *supported***).

5.5 Influence of Dexterity and Age on Speed

Finally, in order to further the understanding of how touchscreen and mouse devices impact performance differently, we wanted to investigate any possible effects of dexterity and age on speed. We note that this investigation was done post-hoc, and so the results should be treated with caution. We divided participants into four equal-size groups, first based on *age*, and second based on levels of *dexterity* according to the sum of their four Pegboard Test scores (as shown in Table 4.1). We used two mixed-design Analysis of Variance (ANOVA) tests, with *device* as the within-subject factor and either *dexterity* or *age* as the between-subject factor. We were only interested in whether touchscreen could minimize the effects of *age* and *dexterity*, thus only interaction effects of device and *age/dexterity* are reported.

For abstract pointing tasks, both higher dexterity and younger age led to faster speed using mouse, but neither factor affected speed on touchscreen: a significant interaction effect between *device* and *dexterity* ($F_{1,15} = 6.21, p = .026, \eta_G^2 = .13$) and between *device* and *age* ($F_{1,15} = 6.29, p = .025, \eta_G^2 = .12$), see Figure 5.3.

For abstract dragging tasks, both younger age and higher dexterity levels led to faster dragging on both devices: but there was no interaction effect of *device* and *dexterity* ($F_{1,15} = .44, p = .73, \eta_G^2 = .03$) or of *device* and *age* ($F_{1,15} = 1.79, p = .20, \eta_G^2 = .03$), see Figure 5.4.

There was no interaction effect between *device* and *dexterity*, nor any interaction between *device* and *age* for any of the C-TOC subtests.

Note that although we refer here to participants with low or high dexterity, all participants were adults experiencing normal aging with no motor deficiencies.

To summarize these findings, touchscreen minimizes the effects of dexterity

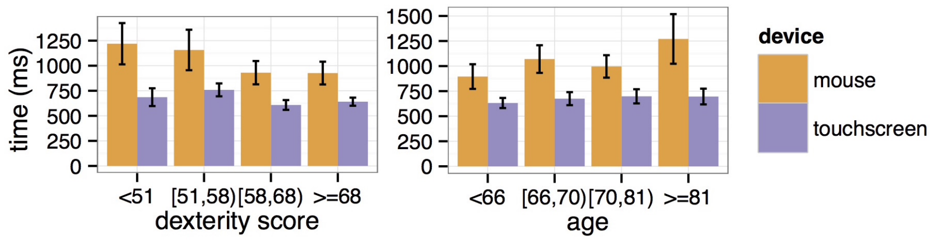

Figure 5.3: Speed for abstract pointing by *device* and *dexterity* (left) or *age* (right).

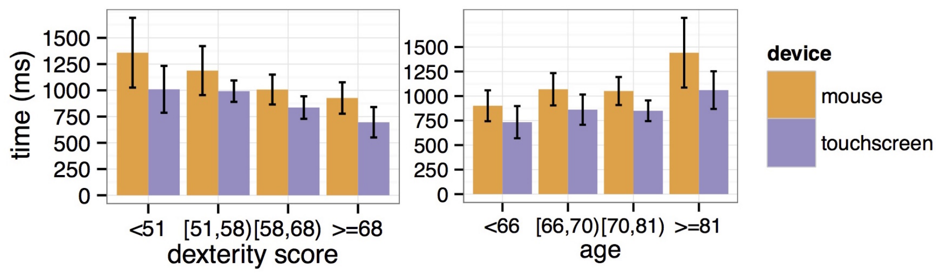

Figure 5.4: Speed for abstract dragging by *device* and *dexterity* (left) or *age* (right).

and age, but only for abstract pointing tasks, not for abstract dragging tasks, nor for any of the real-world C-TOC tests.

Chapter 6

Discussion

We start the discussion with performance results — speed and accuracy — for the abstract and the C-TOC tasks, followed by a KLM-style analysis to explain the speed discrepancy between the two task types. We then look more closely at the different movement strategies participants adopted in the *Pattern Construction* test. We conclude by providing some implications for touchscreen design.

6.1 Abstract Tasks and C-TOC Tasks

Performance results for the abstract pointing tasks are comparable to earlier studies, thus our work replicates those previous findings. In contrast, the dragging results differ from earlier studies: we found older adults are significantly faster doing a dragging task with touchscreen compared to mouse. We suspect the reasons for this difference may be twofold: (1) We only used low-precision levels in the dragging task, and (2) we adopted a variation in the dragging task (the object had to be completely contained in the target region), which is slightly different from a classic dragging task. This may also explain why we had a poorer fit in a Fitts's -style regression model for abstract dragging tasks.

Performance results we obtained for the C-TOC tasks are mostly consistent with those for the corresponding abstract tasks in terms of differences between devices. A touchscreen speed advantage was found for both pointing and dragging on the C-TOC tests, but not all time differences were significant. Accuracy differences between devices were less prevalent with real-world tasks compared to abstract tasks. Lower accuracy on touchscreen was found in only one pointing task, but not in the other three C-TOC tasks.

These results suggest that, with careful calibration, it should be feasible for C-TOC to be self-administered on both touchscreen and mouse-based devices. To determine empirically valid performance calibrations so the devices provide comparable scores, we need a large-scale study with all thirteen C-TOC subtests for speed, accuracy, and other measures that might indicate test-takers' performance, such as number of moves.

6.2 Speed Gain Analysis using KLM

Although many of the real-world tasks show a speed gain on the touchscreen compared to the mouse, there were still some puzzling aspects in the speed results. First of all, not all of the time metrics showed a significant touchscreen advantage, and it was hard to reconcile why a speed gain was observed for one speed metric but not the other. For example, in the two dragging C-TOC tasks, *Sentence Comprehension* had a significant speed gain in the overall task completion time, but none in individual dragging time, whereas the exact reverse was true for *Pattern Construction*. Secondly, speed gain does not translate evenly from abstract to real-world tasks. Some tasks had larger gains than others across devices. Most surprisingly, individual dragging time in *Pattern Construction* had a 53% speed gain[1], which was even higher than in the corresponding abstract dragging task.

It seems that a simple Fitts's law model is not enough to explain the differences between abstract and real-world tasks. The main piece that is missing in a Fitts's model is the cognitive component that always exists (to a varying degree) in any real-world task. C-TOC, a cognitive test, is obviously no exception. In order to better account for both the cognitive and physical components and how the two might interact, we used a Keystroke-Level Model (KLM) [3] to analyze the C-TOC task data.

[1] Time for an individual dragging movement is 83% greater using a mouse compared to a touchscreen. However, as described in Section 5.3, participants make shorter distance moves on touchscreens; the 53% speed gain calibrates for the distance moved.

6.2.1 Action Operators

Four action operators are used in the analysis: *Keystroke* (*K*) is pressing the mouse button or tapping on the touchscreen once the mouse/finger is positioned correctly. *Pointing* (*P*) is pointing to a target. *Dragging* (*D*) is moving an object to a target position while "holding" it throughout. *Mental preparation* (*M*) is the thinking or decision-making involved in doing a task. It is *M* that captures the cognitive parts of the C-TOC tasks. Card, Moran, and Newell's original KLM model [3] kept *M* distinct from the other operators. However, it was obvious from observing the participants that at times they were thinking while positioning an object; i.e., the mental component (*M*) overlapped in time with physical dragging (*D*). We use *DM* to note the cases where participants thought while also dragging.

6.2.2 Operator Sequences

Table 6.1 shows the sequence of operators for each C-TOC task. Both pointing tasks (*Picture-Word Pairs* and *Arithmetic*) have the same operator sequence: participants first derive an answer (*M*), point to the answer button (*P*) and click or tap (*K*) to complete the task. The two dragging tasks (*Sentence Comprehension* and *Pattern Construction*) had largely similar operator sequences that repeat *n* times (*n* being the number of dragging moves) in a trial. For each move, participants start with mental preparation (*M*) where they either recall the shape to acquire (in *Sentence Comprehension*) or they choose a shape and decide where to move it (in *Pattern Construction*). They then point to the shape (*P*), acquire it (*K*), drag the shape to the target position with or without thinking (*D* or *DM*) and then release the shape (*K*). The primary difference in the sequences is in the dragging part – whether the drag is a *D* or a *DM* – which we will argue makes a difference in how the two devices affect performance for the dragging tasks.

6.2.3 Assumptions

We made a few assumptions for the analysis. First, we used the Fitts's law results for abstract tasks to determine the times for actions *P* or *D*, which means doing a *P* or *D* on touchscreen requires less time than doing a *P* or *D* using a mouse. Second, *K* is much faster than *P*, *D*, or *M*, so the difference for *K* between touchscreen and

Table 6.1: Assumptions and operator sequence for C-TOC tasks

Assumptions
$P_{touch} < P_{mouse}, \quad D_{touch} < D_{mouse}$
$K \ll D, \quad K \ll P$
$D < DM \approx M$
Picture-Word Pairs & Arithmetic (pointing tasks)
$T_{task\ completion} = M + P + K$
Sentence Comprehension (dragging task)
$T_{task\ completion} = \sum_{i=1}^{n} (M_i + P_i + K + DM_i + K)$
$T_{individual\ dragging\ move} = DM$
Pattern Construction (dragging task)
$T_{task\ completion} = \sum_{i=1}^{n} (M_i + P_i + K + (D_i \text{ or } DM_i) + K)$
$T_{individual\ dragging\ move} = D \text{ or } DM$

mouse is not a big contributor to task duration. Third, time to finish a DM would be determined more so by M than D, because thinking is more time consuming, in general, than moving.

6.2.4 Analysis

We use KLM to analyze the speed performance of each C-TOC task.

Pointing Tasks

We first analyze pointing tasks. This clarifies why speed gain does not translate evenly from abstract to real-world tasks. We argue that the greater the mental component (M) required by a C-TOC test, the more M will dominate performance and the less likely there will be an effect of device, because device only impacts P or D, but not M. Given that both C-TOC pointing tasks consist of the same operator sequence, when either of these pointing tasks is performed with a device, the only real difference will be the difference in P, determined by the device. Thus, we should expect that the test involving less cognitive workload (lower M) will show a larger effect of device, because the differences in speed (P) from the devices will not be as masked by M. Indeed, the speed gain of the touchscreen over the mouse for

Picture-Word Pairs (32%), a test with low cognitive workload, almost tripled the speed gain in *Arithmetic Test* (12%), a test with high cognitive workload. For neither of these tests did we observe the doubling in speed performance (107% gain) that we observed in the abstract tasks, presumably because M is always present and in C-TOC tasks.

Dragging Task - *Sentence Comprehension*

Dragging tasks are somewhat more complicated. In *Sentence Comprehension* the key to the performance results for *individual dragging moves* is the observation that participants were often dragging the shape to a target position while thinking about the target position, in other words, the drag was a DM, not a D. Given that DM is dominated by M, not D, the results do not indicate significantly faster performance on the touchscreen than a Fitts's model would predict. For the overall *task completion time*, the experiment found it to be faster on the touchscreen compared to the mouse. This can be explained because P is faster on touchscreen than mouse.

Dragging Task - *Pattern Construction*

For *Pattern Construction*, KLM can explain why we observed an even higher speed gain for individual dragging moves in this real-world task compared to a comparable abstract task. For total *task completion time* we found no effect of device; there seems to be a canceling effect at play, caused by an increased number of sub-moves for the touchscreen. On average, each drag (D) was shorter in distance on the touchscreen compared to with a mouse, but there were a greater number of drags (higher n); the two canceled each other out, resulting in no difference in total time. For the other time measure, *individual dragging moves*, we did see a large effect of device. The crux of this is that participants seemed to overlap their thinking with their dragging while using a mouse (i.e., a DM), but they seemed to separate the two much more with the touchscreen (i.e., a D). This is likely because of occlusion – it is harder on a touchscreen to hold a shape and move it while at the same time trying to figure out where to place it because the canvas is partially blocked by one's fingers and hand. Given that dragging on touchscreen is faster than dragging with a mouse ($D_{touch} < D_{mouse}$), a speed gain of 50% the abstract task, and dragging with a mouse is faster than thinking while dragging with a mouse ($D_{mouse} < DM_{mouse}$), by transitivity the speed gain observed between D_{touch} and DM_{mouse} (53%) is higher than the speed gain in an abstract dragging task (46%).

6.2.5 Summary

We have shown that abstract task performance can help to explain performance in a real world task, but it is insufficient on its own. We need to understand how the cognitive component factors into a task and more importantly how it overlaps with other components. KLM is a useful tool in this regard. Of critical importance for the comparison of task performance across touchscreen and mouse was understanding how the two devices differentially impact the overlap between cognition and movement.

6.3 Pattern Construction Task on Touchscreen

There were different behavioral patterns between touchscreens and mouse for *Pattern Construction*. On the touchscreen, participants performed twice as many dragging movements as they did with a mouse. Dragging distance on a touchscreen was relatively shorter, and participants were more likely to decompose a single target-oriented move into smaller consecutive moves (*sub-moves*).

One reason device type contributes to a difference in movement patterns is a bigger overhead in using a mouse during dragging. This comes from two sources: (1) extra workload in pressing down with a mouse compared to just a finger for a single click, and (2) longer time for re-acquiring a shape. Re-acquiring a shape is similar to a pointing task – it is much slower using a mouse than using a touchscreen, according to the pointing task results. The larger overhead of a mouse discourages users from dropping a shape during dragging and then re-acquiring it. One extreme case is the *trial & error* move, which is a single dragging movement that moves a shape to more than one target position. Data showed that *trial & error* moves happen often with a mouse: participants, once having acquired a shape, did not release the mouse button until the shape reached a final target position, often after two or more attempts. We saw much less of this behavior on the touchscreen.

Another reason for the substantial number of moves on the touchscreen was the *sub-move* strategy, which makes dragging tasks much easier. By separating a single target move into n steps of *sub-moves*, task precision for each move is significantly reduced: the distance for each *sub-move* is arbitrarily decided, with an average of A/n (A being the total distance for the move). Dragging on a touchscreen has relatively little overhead, compared to a mouse, and thus there is no extra cost if a single dragging move is decomposed into multiple movements with shorter distance.

Despite the dramatic difference in the total number of moves between devices, no participant seemed aware of the difference. Most participants reported that they "don't think [they] have made more or less moves in either device." Some participants even felt they had more moves using a mouse. Most, but not all, reported having no recollection of making *sub-moves*, even after we demonstrated it. The discrepancy between what participants did and what they believed they did might be due to cognitive chunking when registering *sub-moves*. Users might implicitly chunk all decomposed *sub-moves* into a single target-oriented move. More research is needed to better understand the mechanism of *sub-moves* and the cognitive chunking behind it.

6.4 Implications of Touchscreen Interface Design

We list four implications for touchscreens from the study.

Utilize pointing, but not dragging, to better support aging group. Interfaces that are designed for users with lower dexterity or older age, could best take advantage of touchscreens by adopting more pointing gestures and fewer dragging gestures, to help intended users find easier and faster interaction experience.

Simpler interface: have bigger buttons with reasonable spacing. Making buttons bigger than 40px would help to get rid of the accuracy discrepancy between the touchscreen and mouse. Furthermore, due to smaller touchscreen screen sizes compared to the screens used with a mouse in most commercial products, older adults are more likely to be overwhelmed by cluttered interfaces on a touchscreen, a lesson learned from the not-so-well-designed keyboard in the *Arithmetic* test.

Provide support for decomposable dragging tasks. The *sub-move* strategy indicates a natural tendency to drag differently with a touchscreen than with a mouse. Touchscreen interfaces should support decomposable dragging by allowing objects to "hang" instead of going all the way back to the starting point if users prefer to use the *sub-move* strategy. For example, when a file is dragged into a folder, users could "pause" in the middle, without the file snapping back to its original position.

Explicit usage instruction for capacitive touchscreen to harness the power of touchscreen sensitivity. This experiment used a capacitive touchscreen — iPad 4th-generation. During the pilot, we noticed that older adults still treated the iPad as if it was a resistive touchscreen: they pressed hard to successfully acquire objects. More importantly, most of them, even some that have iPads at home, were not aware that capacitive touchscreens, like the iPad, depend on the conductive nature of human body. They sometimes attempted to use their finger nail to point to smaller objects, and later blamed the insensitivity of the screen. During the actual experiment, all participants were instructed to use their finger tip, not finger nail, to point or drag on a touchscreen. Participants reported that they found "the instruction is especially useful," and the touchscreen was "easier to use" after hearing the instruction.

Chapter 7

Conclusion and Future Directions

Our experiment revealed that a speed gain for touchscreen compared to mouse largely persists from abstract tasks to real-world tasks. Touchscreen performance was more than twice as efficient as using a mouse in abstract pointing tasks, and 46% more efficient in abstract dragging tasks. Though all of our real-world C-TOC tasks demonstrated a speed gain for the touchscreen over mouse (on at least one of our time measures), this could not be explained solely by the abstract task results. A KLM-style analysis better explained the speed gain for our real-world tasks by including a cognitive component that is not required for abstract tasks. We also found that touchscreens yield high error rates on abstract tasks, but less so for real-world tasks. Further research on both Fitts's and KLM-based models to capture the physical and cognitive components of real-world tasks is required.

We found that older adults naturally adopted different movement patterns between devices in one of our C-TOC tasks: on touchscreen, they decomposed a single dragging movement into multiple movements, resulting in shorter individual moves but a greater number of moves compared to mouse. Future work to automate the coding process for identifying movement categorization might be useful in the assessment of cognitive levels and could eventually be integrated into C-TOC.

Our work provides important insights for the C-TOC project. The device-specific difference in performance by older adults on C-TOC tasks suggests a need for a large-scale study to find valid performance calibrations across devices.

Bibliography

[1] R. Bakeman. Recommended effect size statistics for repeated measures designs. *Behavior research methods*, 37(3):379–384, 2005. → pages 18

[2] N. Caprani, N. E. O'Connor, and C. Gurrin. Touch screens for the older user. In *Assistive Technology*, pages 95–118. InTech, 2012. → pages 1, 10

[3] S. K. Card and A. Moran, Thomas P.and Newell. *The psychology of human-computer interaction*. L. Erlbaum Associates Inc., 1983. → pages 5, 32, 33

[4] N. Charness, P. Holley, J. Feddon, and T. Jastrzembski. Light pen versus mouse for a menu selection task: Age, hand, and practice effects. *Human Factors*, 46:373–384, 2004. → pages 5

[5] A. Cockburn, D. Ahlström, and C. Gutwin. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies*, 70(3):218–233, 2012. → pages 5

[6] J. Desrosiers, R. Hébert, G. Bravo, and E. Dutil. The purdue pegboard test: normative data for people aged 60 and over. *Disability and rehabilitation*, 17 (5):217–224, 1995. → pages 15

[7] G. M. Doniger, D. M. Zucker, A. Schweiger, T. Dwolatzky, H. Chertkow, H. Crystal, and E. S. Simon. Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing. *Current Alzheimer Research*, 2(2):117–124, 2005. → pages 7

[8] L. Findlater, J. E. Froehlich, K. Fattal, J. O. Wobbrock, and T. Dastyar. Age-related differences in performance with touchscreens compared to traditional mouse input. In *Proceedings of the SIGCHI Conference on*

Human Factors in Computing Systems, CHI '13, pages 343–346, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1899-0. doi:10.1145/2470654.2470703. URL http://doi.acm.org/10.1145/2470654.2470703. → pages 1, 5, 6

[9] C. Forlines, D. Wigdor, C. Shen, and R. Balakrishnan. Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 647–656, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-593-9. doi:10.1145/1240624.1240726. URL http://doi.acm.org/10.1145/1240624.1240726. → pages 5

[10] ISO:9241-400. Ergonomics of human-system interaction—part 400: Principles and requirements for physical input devices. ISO 9241-400:2007(E), International Organization for Standardization, Geneva, Switzerland, 2007. → pages 11, 13

[11] H. Iwase and A. Murata. Empirical study on the improvement of the usability of a touch panel for the elderly–comparison of usability between a touch panel and a mouse. *IEICE TRANSACTIONS on Information and Systems*, 86(6):1134–1138, 2003. → pages 6

[12] Z. X. Jin, T. Plocher, and L. Kiff. Touch screen user interfaces for older adults: button size and spacing. In *Universal Acess in Human Computer Interaction. Coping with Civersity*, pages 933–941. Springer, 2007. → pages 6

[13] M. Kobayashi, A. Hiyama, T. Miura, C. Asakawa, M. Hirose, and T. Ifukube. Elderly user evaluation of mobile touchscreen interactions. In *Human-Computer Interaction–INTERACT 2011*, pages 83–99. Springer, 2011. → pages 1, 5, 6, 12

[14] M. D. Lezak, D. B. Howieson, E. D. Bigler, and D. Tranel. *Neuropsychological Assessment*. Oxford University Press, 2012. → pages 6, 7

[15] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Hum.-Comput. Interact.*, 7(1):91–139, Mar. 1992. ISSN 0737-0024. doi:10.1207/s15327051hci0701_3. URL http://dx.doi.org/10.1207/s15327051hci0701_3. → pages 8, 18

[16] A. Murata. Eye-gaze input versus mouse: Cursor control as a function of age. *International Journal of Human-Computer Interaction*, 21(1):1–14, 2006. → pages 5

[17] H.-C. Ng, D. Tao, and C. K. L. Or. Age differences in computer input device use: A comparison of touchscreen, trackball, and mouse. In *Advances in Information Systems and Technologies*, pages 1015–1024. Springer, 2013. → pages 1, 5, 6

[18] A. M. Piper, R. Campbell, and J. D. Hollan. Exploring the accessibility and appeal of surface computing for older adult health care support. In *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 907–916, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9. doi:http://doi.acm.org/10.1145/1753326.1753461. → pages 2

[19] T. W. Robbins, M. James, A. M. Owen, B. J. Sahakian, L. McInnes, and P. Rabbitt. Cambridge neuropsychological test automated battery (cantab): a factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders*, 5(5):266–281, 1994. → pages 7

[20] W. A. Rogers, A. D. Fisk, A. C. McLaughlin, and R. Pak. Touch a screen or turn a knob: Choosing the best device for the job. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(2):271–288, 2005. → pages 5, 6

[21] F. Sasangohar, I. S. MacKenzie, and S. D. Scott. Evaluation of mouse and touch input for a tabletop display using fitts' reciprocal tapping task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 53, pages 839–843. SAGE Publications, 2009. → pages 5

[22] N. Schneider, J. Wilkes, M. Grandt, and C. M. Schlick. Investigation of input devices for the age-differentiated design of human-computer interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 144–148. SAGE Publications, 2008. → pages 5

[23] P. J. Snyder, C. E. Jackson, R. C. Petersen, A. S. Khachaturian, J. Kaye, M. S. Albert, and S. Weintraub. Assessment of cognition in mild cognitive impairment: a comparative study. *Alzheimer's & Dementia*, 7(3):338–355, 2011. → pages 7

[24] R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts law research in hci. *International journal of human-computer studies*, 61(6):751–789, 2004. → pages viii, 12

[25] I. Technology. Global touch-screen panel shipments to double by 2016, ihs analyst announces at sid, 2013. https://technology.ihs.com/435487/global-touch-screen-panel-shipments-to-double-by-2016-ihs-analyst-announces-at-sid. → pages 1

[26] J. C. Thompson, C. L. Stopford, J. S. Snowden, and D. Neary. Qualitative neuropsychological performance characteristics in frontotemporal dementia and alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 76(7):920–927, 2005. → pages 7

[27] J. B. Tornatore, E. Hill, J. A. Laboff, and M. E. McGann. Self-administered screening for mild cognitive impairment: initial validation of a computerized test battery. *The Journal of neuropsychiatry and clinical neurosciences*, 17 (1):98–105, 2005. → pages 7

[28] N. Wagner, K. Hassanein, and M. Head. Computer use by older adults: A multi-disciplinary review. *Computers in human behavior*, 26(5):870–882, 2010. → pages 1

[29] K. Wild, D. Howieson, F. Webbe, A. Seelye, and J. Kaye. Status of computerized cognitive testing in aging: a systematic review. *Alzheimer's & Dementia*, 4(6):428–437, 2008. → pages 7

[30] E. Wood, T. Willoughby, A. Rushing, L. Bechtel, and J. Gilbert. Use of computer input devices by older adults. *Journal of Applied Gerontology*, 24 (5):419–438, 2005. → pages 1, 5, 6

[31] S. Zygouris and M. Tsolaki. Computerized cognitive testing for older adults a review. *American journal of Alzheimer's disease and other dementias*, 30 (1):13–28, 2015. → pages 7

Appendix A

Experiment Resources

This appendix contains resources used in the experiment.

A.1 Recruitment Poster

The following study recruitment poster was posted throughout the community. Locations included the UBC campus, Vancouver Public Library branches, Vancouver community and senior centers, and senio housing complexes.

UBC a place of mind THE UNIVERSITY OF BRITISH COLUMBIA

Department of Computer

Usability Evaluation of an Online Cognitive Health Assessment Tool

Study Recruitment

Principal Investigator	**Claudia Jacova, PhD (Medicine)**	
Co-Investigators	Ging-Yuek Robin Hsiung, MD, MHSc, FRCPC	XXX-XXX-XXXX
	Lynn Beattie, MD, FRCPC	XXX-XXX-XXXX
	Philip Lee, MD, FRCPC	XXX-XXX-XXXX
	Dean Foti, MD, FRCPC	XXX-XXX-XXXX
	Sherri Hayden, PhD, R.Psych	XXX-XXX-XXXX
	Joanna McGrenere, PhD	XXX-XXX-XXXX

Purpose This study is designed to investigate how people interact with an online cognitive health assessment tool which involves recall from memory and other cognitive processes. The purpose of this study is to evaluate the usability of the tool's components in order to improve its design.

Participants We are looking for participants aged 55+, who:

- Are healthy, and have normal or corrected-to-normal eyesight, and
- Free of diagnosed cognitive impairments or motor impairments to their hands.

Procedure You will be asked to perform a number of tasks while we record aspects of your performance, including task completion time and response accuracy. You will also be asked interview questions about your experience in performing the tasks, e.g., difficulties encountered. Photographs/Videos may be taken with your permission.

Objective The research objective is to inform and refine the design of an online tool that is intended for cognitive health care purposes. To achieve this, we need to identify any usability issues associated with the tasks to be performed during use of the tool. With this greater understanding, we can continue to design effective and usable health care technologies.

Commitment Your participation in this study will involve 1 session that will require no more than 2 hours of your time and you will be offered a small token or gift for your time.

To Participate Please contact Kailun at XXXXX@cs.ubc.ca or XXX-XXX-XXXX for more information.

A.2 Participant Consent Form

The following is a copy of the consent form participants were required to sign in order to participate in the study. Whenever possible, participants were emailed a PDF copy of the form three days prior to their scheduled session.

The UNIVERSITY OF BRITISH COLUMBIA
Department of Computer Science[1] / Medicine[2]
University of British Columbia
Vancouver, BC, V6T 1Z4

Consent Form

Research Project Title: Development of a Computer-Based Screening Test to Support Evaluation of Cognitive Impairment and Dementia
(Part 1C - Usability Evaluation of an Online Cognitive Assessment Tool)

Principal investigator: Claudia Jacova, PhD, XXX-XXX-XXXX (Medicine)

Co-Investigators:
Kailun Zhang, MSc Student, XXX-XXX-XXXX
Matthew Brehmer, MSc Student, XXX-XXX-XXXX
Joanna McGrenere, PhD, XXX-XXX-XXXX
James Riggs, BSc, XXX-XXX-XXXX
Ging-Yuek Robin Hsiung, MD, MHSc, FRCPC, XXX-XXX-XXXX
Lynn Beattie, MD, FRCPC, XXX-XXX-XXXX
Philip Lee, MD, FRCPC, XXX-XXX-XXXX
Dean Foti, MD, FRCPC, XXX-XXX-XXXX
Sherri Hayden, PhD, R.Psych, XXX-XXX-XXXX
Sung-Hee Kim, PhD, XXX-XXX-XXXX

In this study, we aim to identify usability issues associated with selected task components of a novel computer-based cognitive test battery, called Cognitive Testing on Computer (C-TOC). You are being invited to participate in this study because you are 55 years of age or older with or without any diagnosed cognitive impairments or motor impairments to your hands. Your participation will help us probe the usability of C-TOC task components.

Your participation in this research study is entirely voluntary. This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

If you wish to participate, you will be invited to sign this form but you should understand that you are free to withdraw your consent at any time and without giving any reasons for your decision.

Purpose: This study is designed to investigate how people interact with an online cognitive health assessment tool which involves recall from memory and other cognitive processes. The purpose of this study is to evaluate the usability of the tool and improve its design.

Procedure: Your participation in this study will involve 1 session that will require no more than 2 hours of your time. During this session, you will be asked to perform a number of tasks on a desktop computer. We will record aspects of your performance, including task completion time and accuracy. This test is not meant to test your skills or experience with computers; it is only being carried out to probe the usability of C-TOC task components. You will also be asked interview questions about your experience in performing the tasks, e.g. difficulties encountered. In all circumstances, you do not need to answer any questions that you do not feel comfortable answering.

Objective: The research objective is to inform and refine the design of an online tool that is intended for cognitive health care purposes. To achieve this, we need to identify all usability issues that may affect people's performance on the tasks that are presented in the online tool. This knowledge will help us design effective and usable health care technologies.

Option for Photographing/Videotaping:
For the purpose of data analysis, we would like to videotape your computer session and your interview. Please note that this is an optional procedure, which you are free to decline, and a refusal to videotape will in no way affect your eligibility for this study. Only the investigators of this study will have access to the recordings. The recordings will be stored in a secured departmental network of Neurology for three years after the study, which will then be permanently erased. Participants' identity will be protected by masking in publications and presentations. Please check and initial the ones you agree.

- I agree that the researchers may videotape my computer session. __________
- I agree that the researchers may videotape my interview. __________

What are the Possible Harms and Side Effects of Participating?
You may experience fatigue from performing the computer tasks and answering the questions.

What are the Benefits of Participating in this Research?
There may be no immediate, direct benefit to you as a result of participating in this study. However the findings from this study can help us improve future health care technologies that may benefit you, your family members and the community in the longer term.

What Happens If I Decide to Withdraw My Consent to Participate?
Your participation in this research is entirely voluntary. You may withdraw from this study at any time, and are not required to provide any reason for withdrawing. If you choose to enter the study and then decide to withdraw at a later time, all data collected about you during your enrollment in the study will be retained for analysis. By law, this data cannot be destroyed. If you wish to withdraw your consent, we ask that you notify Kailun Zhang at XXX-XXX-XXXX, or James Riggs at XXX-XXX-XXXX.

What Happens If Something Goes Wrong?
Signing this consent form in no way limits your legal rights against the study sponsor, investigators, or anyone else.

Will My Taking Part in this Study be Kept Confidential?
Your confidentiality will be respected. The Investigators in this study will be responsible for maintaining your confidentiality at all times. Study records will be labeled only with an assigned numeric code. They will not include information that identifies you by name, initials, or date of birth. This code number and the connection of the code number to your name and identifying information will be stored in a private, password-protected computer in the Department of

Neurology at UBC Hospital. Access to personal identifying information will be restricted to the Principal Investigator, Co-Investigators, and their research study staff.

Results from this study may be presented at meetings and may be published, but no information that discloses your identity will be released or published without your specific consent to the disclosure. However, research records and medical records identifying you may be inspected in the presence of the Investigator or his or her designate, and the UBC Research Ethics Board for the purpose of monitoring the research. However, no records which identify you by name or initials will be allowed to leave the Investigators' offices.

Who do I Contact if I have any Questions or Concerns about the Study?

If you have any questions or desire further information with respect to this research, you should contact Dr. Joanna McGrenere at XXX-XXX-XXXX or Kailun Zhang at XXX-XXX-XXXX. If you have any concerns about your rights as a research subject and/or your experiences while participating in this study, you should contact the Research Subject Information Line at the University of British Columbia's Office of Research Services at XXX-XXX-XXXX.

Subject Consent to Participate:

- I have read and understood the subject information and consent form.
- I have had sufficient time to consider the information provided and to ask for advice if necessary.
- I have had the opportunity to ask questions and have had satisfactory responses to my questions.
- I understand that all of the information collected will be kept confidential and that the results will only be used for scientific objectives such as research and publications.
- I understand that I can refuse to answer any questions that I do not feel comfortable answering from this study.
- I understand that my participation in this study is voluntary and that I am completely free to refuse to participate or to withdraw from this study at any time.
- I understand that I am not waiving any of my legal rights as a result of signing this consent form.
- I understand that there is no guarantee that this study will provide any benefits to me.
- I have read this form and I freely consent to participate in this study.
- I have been told that I will receive a dated and signed copy of this form.

Signatures

Printed Name of Participant Signature and Date

Principal Investigator or designated representative Signature and Date

A.3 Questionnaires

A.3.1 Demographics

Participants were asked to complete a short demographic questionnaire after signing the consent form.

Demographic

1. **participant number**
   for experimenter to fill in

   ....

2. **Age**

   ....

3. **Gender**
   Mark only one oval.

   ◯ Female
   ◯ Male
   ◯ Prefer not to say

4. **Handedness**
   Mark only one oval.

   ◯ Right
   ◯ Left
   ◯ Ambidextrous

5. **First/Dominant Language**

   ....

6. **How would you rate your vision (with corrective lenses, if required)**
   Mark only one oval.

   ◯ Excellent
   ◯ Good
   ◯ Fair
   ◯ Poor

7. **Do you experience any colour deficiency?**

   ....

8. **Medical conditions that affect motor functions?**

   ....

9. **How often do you use ...**
   Mark only one oval per row.

	Less than once per week	Once or twice per week	Several times per week	Once or twice per day	Several times per day
Computer w/ mouse	◯	◯	◯	◯	◯
Touch based tablet or phone	◯	◯	◯	◯	◯
Digital devices in general	◯	◯	◯	◯	◯

10. **Which of the following applications do you use regularly (at least once a week)?**
    Tick all that apply.

- [ ] Word processor (like Microsoft Word)
- [ ] Web browser
- [ ] Email
- [ ] Games
- [ ] Media players (like Windows Media Player, QuickTime)
- [ ] Other: ....

Powered by
Google Forms

A.3.2 Interview Scripts

The following interview script was used at the end of the experiment session.

Interview Questions

First we will go through each test you have done and **make comparisons between the test on a computer w/ mouse and on a touch-based tablet.**

Go through C-TOC test one by one

1. [comparison across devices] For this test, do you prefer to do it on a computer with a mouse or on a touch-based tablet?
2. [comparison across devices] Do you have any difficulties in doing the cognitive test on a computer with mouse/touch-based tablet?

Show both abstract tasks

1. For the pointing task, which is harder: using a mouse or using touch? Explain.
2. For the dragging task, which is harder: using a mouse or using touch? Explain.

Now we are going to make comparison between tests.

1. Both Picture Word Pair and Arithmetic are tests require pointing. Did you notice any difference between the two?
2. [If yes for the previous question] How is the pointing task different between the two tests? Would this difference (in difficulty level) have any impact on your preference to do the two tests on touch-based tablet versus computer w/ a mouse?
3. Both Sentence comprehension and pattern construction are tests require dragging Have you noticed any difference between the two?
4. [If yes for the previous question] How is the dragging task different between the two tests? Would this difference (in difficulty level) have any impact on your preference to do the two tests on touch-based tablet versus computer w/ a mouse?

For Pattern Construction Tasks

We are now going to discuss the pattern construction test in more detail, as it involved a lot of interactions on both devices.

1. In terms of the number of moves that you made (the total number of times you moved any of the shapes), do you have a sense of whether you made more, less, or about the same number of moves using the touch screen compared to the mouse? Why
2. Did you ever attempt to rotate any of the shapes? Do you recall what device (touch or mouse) you were using? Why did you attempt to rotate or why not?
3. Sometimes shapes were in the way of where you needed to place other shapes in order to construct your pattern. What strategy did you to deal with shapes that were in the way? <If they don't offer anything, suggest: allow shapes to overlap in order to minimizes number of moves and moving the shapes out of the way to clean up the canvas>
4. Was there one device (touch screen or mouse) on which that you felt compelled to do make your moves more precise?
5. Do you ever separate one move of a single shape (to get to a particular destination) into multiple smaller sub-moves (to get to that same destination) (on a touch screen)? Why?

Experiment related question

- Have you experienced fatigue during the experiment? If yes, when does it start?

General questions

1. How familiar are you with computer with a mouse or a touch-based tablet? (have you done dragging on tablet before?)

2. Do you have a computer with a mouse in your home?

3. For what purposes do you use it? (work, leisure, or household purposes)

4. Do you have a touch-based tablet in your home?

5. For what purposes do you use it?(work, leisure, or household purposes)