HOW DO TACTILE INTERACTIONS AND MOVEMENT ORIENTATIONS AFFECT INFORMATION PROCESSING OF SEARCH ENGINE RESULTS PAGES (SERPS) ON MOBILE DEVICES?

BY

YILIN REN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Communications & Media
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Kevin Wise, Chair
Professor Alejandro Lleras
Associate Professor Brittany Duff
Professor Mike Yao

ABSTRACT

Interaction between humans and media has become increasingly tactile and mobile. Users interact with touchscreens frequently to consume information, in part because the screen size of mobile devices is much smaller than that of computers. More specifically, people scroll upward to view additional content or swipe horizontally to view the next page. There are two intersecting aspects of this phenomenon: Tactile interaction—scrolling versus swiping, and movement orientation—horizontal versus vertical. Scrolling navigates continuously displayed information—gradually replacing old with new, not unlike unrolling a traditional scroll. Swiping typically involves more discrete transitions—new information completely replaces old information, much like turning a page in a book. These simple interactions dictate virtually all of our consumption of media on mobile devices.

Though some previous studies explored the effect of tactile interactions on users' attention and memory, they failed to consider the effect of movement orientation inherent in the interactions. Two studies were conducted to explore how tactile interaction and movement orientation, in both isolation and combination, affect individuals' search performances and user experiences.

Study 1 included 276 participants recruited from Amazon M-Turk and Study 2 included 172 undergraduate student participants from a large Midwestern university. All participants used their smartphone to perform a series of search tasks while using one of four possible interaction-orientation combinations: vertical scrolling, horizontal scrolling, vertical swiping, and horizontal swiping. After participants finished the search tasks, their self-reported time perception, power usage, and user satisfaction were measured. Mix-effects modelling was used to analyze both main and interaction effects of tactile interaction and movement orientation on the aforementioned dependent variables. The results showed a significant main effect of tactile interaction, in which users performed better when using scrolling interface than swiping interfaces. More specifically, while the search accuracy was the same, participants searched faster when using scrolling interface comparing with swiping interfaces. While the search time was the same, participants achieved higher search accuracy while using a scrolling interface compared to a swiping interface. The implication of this study is that it teased apart tactile interaction from movement orientation to investigate both the individual effect and interaction effect on users' task performance and user experiences, while previous studies mixed these two factors together. In addition, it showed that the type of information (picture vs. text) might interact with different cognitive load placed by different interfaces and generate different patterns of users' task performance data. Practically, when designers consider different tactile interactions for a visual search interface, scrolling mode might be better than swiping mode in terms of generating better search performance while achieving the same level of user satisfaction.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the longstanding support of many people. I would like to express my sincere gratitude to my advisor Dr. Kevin Wise for his guidance, patience, and support through my doctoral study. You were always there when I made changes about my dissertation topics, when I switched my career paths, and when I had hard times in life. I greatly appreciate your help and encouragement during this journey. I would also like to thank my committee members, Dr. Alejandro Lleras, Dr. Brittany Duff, and Dr. Mike Yao, who discussed research ideas and studies with me and gave me insightful comments. Finally, thanks to my parents and friends who always offer support and love.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Introduction

When users consume media content on mobile devices, they use various techniques to navigate and interact with the interface and content. For instance, users can drag their fingers across the screen to orient themselves when using an online map. As another example, they could use their fingers to spin a product image to view it from different angles on an e-commerce website. Different human-device interactions provide diverse user experiences and influence the way users process media content, such as enjoyment level or learning outcomes.

While many scholars use the phrase "interaction techniques" to refer to these human-device interactions, I use a more specific term, "tactile interaction." The former phrase could include oral/aural interaction (as with Siri or Alexa), while my terminology is limited to interaction techniques that specifically rely on touch. Tactile interaction includes scrolling, swiping, tapping, zooming in or out and so on; my study focuses solely on scrolling and swiping. Scrolling navigates continuously displayed information—gradually replacing old with new, not unlike unrolling a traditional scroll. Swiping interacts with information that is "discretely" displayed—new information completely replaces old information, much like turning a page in a book. The current study investigates these two tactile interactions as they intersect with movement orientation—horizontal and vertical. These intersections provide four interaction combinations: horizontal swiping, horizontal scrolling, vertical swiping, and vertical scrolling. (see Table 1).

Previously, two of these combinations were almost exclusively used in interface design: horizontal swiping (e.g., eReader or weather app) and vertical scrolling (e.g., traditional websites and Google search results), both starred in Table 1. More recently, mobile device interfaces have begun to use the other two as well. For example, some ads and multimedia (e.g., videos or images) appear in Google search results as a featured section along a continuous band, allowing users to scroll horizontally for further viewing. Additionally, in prominent video apps like TikTok, users navigate through different videos by vertical swiping.

Table 1

Four Different Human-Device Interactions and Examples

		Movement Orientation	
Tactile Interaction		Horizontal	Vertical
	Swiping (discrete presentation)	**eReader and weather apps** ★	TikTok / some video apps
	Scrolling (continuous presentation)	google search featured results sections	**Normal Google / webpage** ★

1.2 Background of the Topic

Americans spend about 5.4 hours using their phones each day (Milenkovic, 2020). Many activities that we used to do on a computer are now done using mobile devices (e.g., sending emails, searching for information, or reading news). Since mobile device screens are typically much smaller than computer screens, users interact more frequently with the touch face when navigating through media content. Among various mobile device interactions, scrolling and swiping are the most common. Big tech companies like Google and Facebook keep developing and launching new combinations of mobile device interactions to enhance user experiences. It is important to know how these interactions influence users toward specific goals such as searching, shopping or entertaining.

Some previous research has explored how scrolling versus swiping influences users' performance and experience of a task, such as mobile app shopping and online searching. For example, Choi, Kirshner, and Wu (2016) compared swiping horizontally versus scrolling vertically in a shopping interface. They found that swiping leads to higher cognitive absorption and playfulness than scrolling, which suggests that shopping apps with a swiping interface may be more successful than others in this context. Kim et al. (2016) compared users' search performance while using a pagination (similar to swiping) or scrolling interface on a mobile phone, and found participants had better search speed and accuracy when using pagination. These results may prompt search-engine designers to present search results in discrete pages to enhance users' search performance.

However, these studies (and others) have limitations in terms of asymmetrical comparisons. For example, in the mobile app shopping study above (Choi et al., 2016), they compared a one-step process (swipe right to save, swipe left to dismiss) with a two-step process (scroll and tap), rather than a pure comparison of swiping versus scrolling. In the search engine study (Kim et al., 2016), they compared horizontal swiping versus vertical scrolling, thus making it difficult to tease apart the interplay between tactile interactions and movement orientation.

My proposed study would employ symmetrical comparisons to test the effect of tactile interaction and movement orientation combinations on users' processing of media content.

1.3 Purpose of the Study

The goal of this study is to inform future mobile device interface design, based on investigating how the combinations of tactile interactions and movement orientations impact information processing of media content—especially for search engine result pages (SERPs) that display multiple products. There are three objectives that will lead to achieving this goal. The first two objectives are to investigate the (1) individual and (2) combined effects of tactile interaction and movement orientation on how users process media contents. The third objective is to assess users' subjective experiences when using different combinations of tactile interaction and movement orientation.

CHAPTER 2: REVIEW OF THE LITERATURE

The existing literature suggests that users' experiences might differ while consuming media content using two different tactile interactions (swipe vs scroll). Scrolling provides a seamless experience in that additional content is revealed together with some of the previous content, until the end of the page. Also, for a specific piece of information, it only has a relative position rather than an absolute position because each scroll changes its position. Conversely, every swipe gesture reveals a page of brand-new information that has no overlap with the previous page. Also, each piece of information has its absolute location. Having an absolute versus relative location will influence the cognitive load on users because spatial information is encoded while processing information (Piolat, Roussey & Thunin, 1997).

This chapter first addresses relevant literature on cognitive load, as it is the primary mechanism that impacts information processing. I will then address human-computer interactions under the broad consideration of modality interactivity. The rest of the literature review will home in on tactile interactions (one kind of modality interactivity) on touch screen devices—alone and combined with movement orientations—and their effects on cognitive processing of media contents.

2.1 Cognitive Load

2.1.1 The Concept of Cognitive Load

Cognitive load is a prominent concept in psychology, media psychology and educational psychology. The basic definition—holding items temporarily in working memory which taxes one's executive control functioning (Lavie, 2005; Logan, 1978) is consistent across these areas, but existing scholarship in educational psychology is most relevant to the phenomena explored here. Therefore, the following discussion of cognitive load is derived primarily from the educational psychology literature.

Cognitive load is defined as a multidimensional construct representing the load that performing a task imposes on an individuals' cognitive system (Paas & van Merriënboer, 1994). According to Paas & van Merriënboer, cognitive load consists of causal factors that affect cognitive load and assessment factors that are affected by cognitive load. The three causal factors are task characteristics (e.g., structure, type of reward system, time pressure, noise), subjects' characteristics (e.g., cognitive ability, prior knowledge), and interactions between them. The assessment factors include mental load, mental effort, and performance. Mental load comes from a task-centered dimension, which refers to the load imposed by the task. Mental effort comes from a human-centered aspect, which refers to the actual cognitive resources allocated to a specific task. Paas and van Merriënboer suggest that both mental effort and performance reflect all three causal factors (i.e., task characteristics, subjects' characteristics, and interactions between them). Since the intensity of effort spent on tasks is often considered the essence of cognitive load (Hamilton, 1979; Paas, 1992), scholars also use mental effort as an index of cognitive load (Cooper, 1998).

Researchers have conceptualized cognitive load as having three specific types: intrinsic, extraneous, and germane load (Schnotz & Kürschner, 2007; DeLeeuw & Mayer, 2008; De Jong, 2010; Kalyuga, 2011). Intrinsic load refers to how conceptually complex the materials are. Extraneous load (ineffective load) refers to the cognitive load that results from mainly poorly designed instruction. Germane load (effective load) is caused by effortful learning that assists in schema construction (Sweller, 1999; Schnotz & Kürschner, 2007). In short, the intrinsic load is decided by the content, the germane load is determined by individuals' effortful learning, and extraneous load is everything else related to the information acquiring experience. Previous studies tested extraneous load by using different instructional procedures on the same materials (Paas, Renkl & Sweller, 2003; Chandler & Sweller, 1996; Cierniak, Scheiter & Gerjets, 2009), my experiment uses different interfaces for the same materials, as an alternative approach to how these interfaces can also affect extraneous load, and therefore cognitive load.

2.1.2 Measuring Cognitive Load

According to the assessment factors of cognitive load from Paas & van Merriënboer's (1994) model, it is straightforward that cognitive load can be measured by mental load, mental effort, and performance. Some researchers used analytical methods to access mental load through mathematical models and task analysis, together with expert opinion (Sweller, Van Merrienboer & Paas, 1998; Xie & Salvendy, 2000). Other researcher used empirical methods to access mental effort and task performance, mainly through subjective rating scales, primary and secondary task performance, and psychophysiological methods such as heart rate and pupil dilation (Paas, 1992; Paas & van Merriënboer, 1994b; Beatty & Lucero-Wagoner, 2000; Chandler & Sweller, 1996; Brunken, Plass & Leutner; 2003).

In terms of the analytical approach to access mental load, Sweller (1988) adopted a production system to model the difference in working memory load between a forward-working strategy and mean-ends analysis. The conventional means-ends approach means that a problem solver works backward from the goal state and setting sub-goals that can reduce the differences between the goal state and initial problem state by coming up with problem-solving operators. This means that problem solvers continually hold and process the current problem state, the goal state, and the relation between these two in working memory. Research found that most errors happened during the sub-goal rather than the goal phase (Ayres, 2001). A goal-free forward-working strategy eliminates the backward-working phase and begins by choosing operators using what is given to go to the next state until reaching the goal state. Sweller (1988) used computational models that took the number of statements in working memory, the number of productions, the number of cycles to solutions, and the total number of matched conditions into account, and found that using goal free forward-working strategy substantially reduces cognitive load compared with mean-ends strategy.

Subjective rating scales that are used to access mental effort are based on the assumption that individuals can introspect on their cognitive processes and report the amount of mental effort they spend on tasks (Paas, Tuovinen, Tabbers & Van Gerven; 2003). Most subjective measures are multidimensional that groups associated variables such as mental effort, fatigue, and frustration together (for a review, see Nygren, 1991). For example, the NASA task load index (TLX) accesses cognitive load in six aspects: mental demand, physical demand, temporal demand, performance, effort and frustration level (Hart & Staveland, 1988). Few studies used unidimensional scales (Paas & van Merriënboer, 1994b). Participants indicate their experienced cognitive load by selecting a point from semantic differential scales, for example, a scale ranging from 1(very, very low mental effort) to 9 (very, very high mental effort) after they finish each task (Paas, 1992).

Task performance measurement techniques that access cognitive load include two sub-categories: primary task performance (e.g., comprehension accuracy for a reading task) and secondary task reaction time, accuracy- and error- rate. The secondary task paradigm assumes the performance on the secondary task can reflect the cognitive load imposed by the primary task. More specifically, the higher cognitive load imposed by the primary task, the less available cognitive resources would be allocated to the secondary task, resulting in worse secondary task performance. Scholars typically use a simple attention sustained task such as detecting a visual or auditory signal as a secondary task (Lang & Basil, 1998).

Physiological techniques that are used to access cognitive load are based on the assumption that physiological variables can reflect the intensity of individuals' cognitive functioning (Paas et al., 2003). For example, Paas and van Merriënboer (1994b) used heart-rate variability to estimate the amount of cognitive load. However, they found this measure invalid and insensitive to a subtle change in cognitive load. Beatty and Lucero-Wagoner (2000) used task-evoked pupillary responses (TEPRs) method, which accessed mean pupil dilation, peak dilation, and latency to the peak to measure cognitive. This method was shown to be sensitive and used in later studies (e.g., Van Gerven, Paas, van Merriënboer & Schmidt, 2004).

Chen, Epps and Chen (2011) compared four measurements of cognitive load: self-report task difficulty, task completion time, performance accuracy and physiological features (i.e. a combination of pupil size and blink number) and found that self-report is the most sensitive and accurate method. Based on this, I used the subjective rating scale similar to that by Cierniak, Scheiter and Gerjets (2009) and asked participants how difficult they feel about each search task.

The methods and analyses described above measure cognitive load, which allows researchers and developers to make more informed decisions when designing materials, instructions, and interfaces.

2.2 Humans and Computers: Modes of Interactivity & Interactions

Human-computer interaction (HCI) is a multidisciplinary field of study focusing on the interaction between humans and computer technology (Montuschi, Sanna, Lamberti & Parvati; 2014). Though at the beginning it refers to computer technology, today it has expanded to cover almost all forms of information technology design (Yao & Ling, 2020). It is sometimes termed as human–machine interaction (HMI) or man-machine interaction (MMI) (Karray et al., 2008). The primary goal of HCI is to explore design and implementation that improve the interaction between humans and machines (Montuschi et al., 2014).

In terms of the interaction itself, Sundar (2007) proposed an Interactivity Effects Model (IEM), which identifies three types of interactivity: source interactivity, message interactivity, and modality interactivity. Source interactivity means that users can be the source or gatekeeper of information (e.g., customize the interface to present personalized information). Message interactivity occurs when a website allows the exchange of messages between the system and users (e.g., using hyperlinks and buttons to increase message interactivity). Modality (or medium) interactivity refers to the various methods of interaction (e.g., clicking, scrolling) offered by the interface. All three forms of interactivity can impact the user's engagement with content, thus influencing users' cognitive, attitudinal, and behavioral responses to the content. This current study only focuses on the interactions under modality interactivity, as shown in Figure 1.

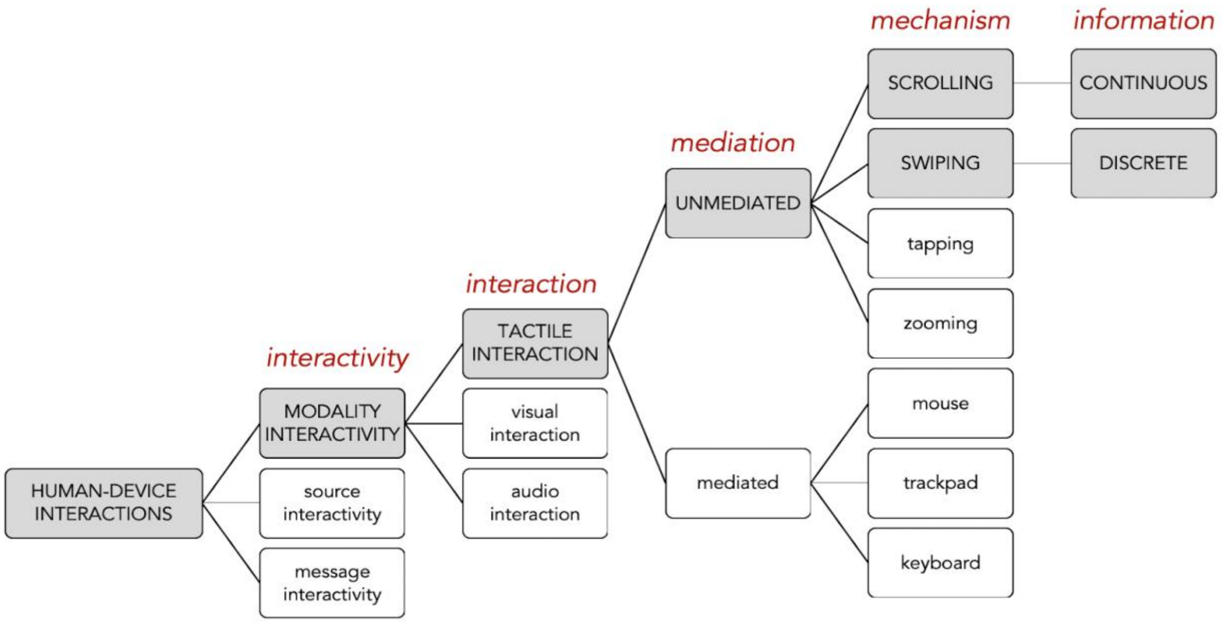


Figure 1. Structure and scope of the prominent tactile interactions with media. Grey shading indicates items of focus in this study.

2.2.1 Modality Interactivity

Modality Interactivity means the variety of modalities that have been offered to users to experience various parts of a website, from simple texts to graphics, animation, audio, and video (Sundar, 2007, p90). Modality interactivity includes oral/aural interactions (as with Siri or Alexa), visual interactions (as with body movement capture by computer camera), and tactile (touch-based) interactions. Different modality interactions—whether by type or number—influence users' experiences and performances of certain tasks.

For example, Xu and Sundar (2014) allowed e-commerce users to rotate, zoom and hover over images of a product rather than just clicking and scrolling. They found that these higher levels of modality interactivity gave users a greater sense of control over the content and made the experience more fun, which led to a more positive attitude toward the website and products.

Oh, Robinson and Lee (2013) tested whether using a page-flipping (i.e., hold and drag the mouse at the corner to flip) versus clicking (i.e., double click the corner) interface to read an online magazine would influence user's attitude towards the website and memory of the magazine content. They found that participants had a worse recall and recognition memory for the content when they used a page-flipping interface, compared with the clicking one. However, participants evaluated the page-flipping website as more usable and engaging, leading to greater behavior intention towards the website.

Oh and Sundar (2015) run another study to test how the presence and absence of a slider on a website influences user's processing of health messages (i.e., anti-smoking). They found that participants reported being more absorbed and less distracted when using the website that has a slider and had a more favorable attitude towards both the website and media message. However, participants in the slider condition generated a lower number of antismoking thoughts, indicating a worse message elaboration, compared with the control condition which does not have a slider. They suggested that modality interactivity can have a positive effect on cognitive absorption, but at the cost of elaboration, because high interactivity might preoccupy the user in experiencing the site, so the storing and reflecting upon the content would be impaired. They also mentioned that according to the limited capacity model of mediated message processing (Lang, 2000), adding more interactive features to the interface can deplete the pool of available cognitive resources, thereby impairing the information processing of the message.

These findings show that high modality interactivity is related to positive perceptual assessments like the ease of use, vivid experiences, and further increased user engagement and attitudes, but sometimes at the cost of cognitive elaboration (e.g. storing and recalling) of the media content. In my current study, I tested different combinations and levels of modality interactivity, in order to assess their influence(s) on users' experiences and task performances.

2.2.2 Tactile Interactions

Tactile interactions are touch-based techniques used to interact with the media interface. These can be performed with or without a separate input device (e.g., keyboard, mouse, trackpad) and include scrolling, swiping, clicking, zooming, hovering, and sliding. When people use the mouse and keyboard to interact with devices like personal computers, it is mediated tactile interaction; when they use their fingers to touch the screen to directly interact with media content, it is unmediated tactile interaction. Though these two differ from each other (as the unmediated provide a more direct/integrated experience of interaction with devices than mediated), they share many similarities. For example, tapping on a screen (unmediated) is similar to using a mouse to click (mediated). The differences between mediated and unmediated tactile interaction is not the scope of this current study.

My study focuses on unmediated tactile interactions—specifically scrolling versus swiping. Choi, Kirshner, and Wu (2016) compared these two on a mobile app shopping task. They found that swiping interface (one product each time, swipe right to save, swipe left to dismiss) leads to higher cognitive absorption and playfulness than scrolling interface. They also found that higher cognitive absorption leads to better performance (the users' ability to save products they like) and reuse intention while playfulness didn't. In their study, task performance was measured by the number of discarded items after participants performing the "viewing and saving to like" task and a questionnaire. A small number of discard items indicate higher task performance.

However, their experimental design obscures any direct relationship between tactile interactions and shopping experiences. For instance, horizontal swiping is a one-step process—users make a single swipe in one direction or another to save or dismiss each item. Vertical scrolling, however, is a two-step process—users first navigate products by scrolling, then tap an item to save it. Furthermore, users are able to revisit “unsaved” items in the scrolling context; but cannot “undismiss” items to view again in the swiping context. In essence, Choi, Kirshner, and Wu (2016) were testing scrolling (as continuous information) versus swiping (as discrete information); the inconsistent number of process-steps, and the two different movement orientations, are all confounding variables that prevent a focused, isolated study of scrolling versus swiping.

Within the broader category of tactile interactions, movement orientation is particularly relevant and a focus of this study. Here, movement orientation refers to the direction of finger movements (horizontal versus vertical) when users interact with a touch screen. Few studies have looked at the effect of movement orientation while controlling tactile interaction. For example, Ren, Kang, Ryu and Han (2017) found that high Need-For-Touch participants prefer horizontal-swipe than vertical-swipe products, while low Need-For-Touch participants didn’t show such difference. Braganza et al. (2009) found that more participants prefer vertical scroll than horizontal scroll, though their reading performance did not differ accordingly. These studies showed that movement orientation can influence user preferences, but generally lack of evidence concerning its influence on task performance. To address this issue, the study proposed here includes four conditions: vertical scrolling, vertical swiping, horizontal scrolling, and horizontal swiping. This provides the ability to test the main and interaction effects of tactile interaction and movement orientation on task performance, in essence teasing them apart as previous studies have not been able to do.

2.3 Information Processing of Media Content

In order to understand information processing of media content, this project isolated relationships between and among cognitive load, tactile interactions, and movement orientations. Below, I first addressed how tactile interactions alone (specifically, scrolling and swiping) affect cognitive load. I considered the same processes with an added variable—movement orientation. This illuminated the complex network of variables in play when users interact with a computer interface.

2.3.1 Tactile Interactions, Cognitive Load, and Information Processing

This section discusses why two types of tactile interaction--swiping versus scrolling-- can influence cognitive load and information processing. According to the Limited Capacity Model (Lang, 2000), consumers of media have a limited pool of cognitive resources available at any given time. Differences in the amount of cognitive resources required by one task influence the distractibility (Lavie, 2010) and available cognitive resources for a secondary task (Lang & Basil, 1998). Besides the objective measurement of task performance and memory, I also discuss how tactile interaction may influence users' relevant subjective experiences, particularly time perception and user satisfaction.

2.3.1.1 Effects of Tactile Interactions on Media Content. Different tactile interactions (here scrolling and swiping) induce different cognitive load because of the amount of modality interactivity inherent in the interaction, as mentioned in the Section 2.2.1. In addition, for specific information (e.g., a word), it has an absolute location in the swiping condition but a relative location in the scrolling condition. Having an absolute versus relative location will influence the cognitive load on users because spatial information is encoded while processing information (Piolat, Roussey & Thunin, 1997).

Previous studies show that positional information is processed during reading, which means that participants assign an address in the form of (x, y) coordinates to each piece of information (Piolat et al., 1997). This concept is referred to as “spatial encoding”, and it can help readers to improve subsequent recall of information in long texts (Lovelace & Southall, 1983); in short, the “where” may serve to recall the “what” (Baccino & Pynte, 1994). When we consume media on scrolling interfaces (inherent in the continuous information display), each piece of information lacks a static "place on the page" location, which might impair the reallocating process if readers try to reinforce surface memory (Sanchez & Wiley, 2009). In other words, each time when users scroll, the spatial layout is disrupted, making it difficult for users to create a two-dimensional representation of the entire information (Piolat et al., 1997), thus increase the cognitive load on users. Some readers might use other reference points like the position of words with respect to other words (Swanston, Wade & Day, 1987) to create a mental spatial representation in the scrolling interface, thus also increasing the cognitive load, compared with the swiping interface (inherent in the discrete display) when each piece of information has its two-dimensional representation.

Piolat et al. (1997) conducted two studies comparing the effects of page-by-page (i.e., discrete) versus scrolling (i.e., continuous) text presentation on users’ performance on reading and revising texts. They found that when presenting short expository texts in discrete pages, participants have a better memory for details about the content, compared with continuous scrolling conditions. Also, they found users experiencing discrete text presentation generated more coherence revisions (to improve the text) than those experiencing continuous text presentation. Sanchez and Wiley (2009) compared users' reading comprehension while using a single continuous scrolling webpage versus 13 discrete pages of the same content (a 2700-word-long text). They found that participants have a significantly better comprehension in the discrete presentation condition compared with continuous scrolling presentation condition. They also found that lower working memory capacity (WMC) individuals perform significantly worse than higher WMC individuals in continuous presentation conditions, but not in discrete presentation conditions.

Kim et al. (2016) did a study to explore whether using different control types (scrolling vs. pagination) to complete a search task (e.g., find the target among ten google search results) on a mobile phone would influence search performance and user experiences. Each participant experienced two control types with six different target positions (12 trials in total). Results show that participants spend longer to identify the correct result when they use scrolling compared with pagination. Also, if the correct answer is located on the second page, pagination results in better search accuracy than scrolling. Wang and Yueh (2014) found that users of all reading levels demonstrate a preference for forms that are provided in a series of chunks as opposed to a long, scrollable page, which indicates that readers may perform better when the information is presented in discrete pages than scrolling.

To summarize, the existing literature shows that positional information is processed during reading. However, in terms of experiment design, previous studies always combined scrolling with vertical movement and discrete presentation with horizontal movement. The current study aims to tease the tactile interaction (scroll vs. swipe) and movement orientation (vertical vs. horizontal) apart to explore their individual effect. Thus, four conditions: vertical swiping, vertical scrolling, horizontal swiping, and horizontal scrolling are created.

As we know, the x and y coordinates of each piece of information on the screen (e.g., page one, line 3) are fixed in swiping (i.e., discrete presentation). But, in terms of scrolling (i.e., continuous presentation), the spatial layout is disrupted each time users scroll. This makes it difficult for users to create a two-dimensional representation of the entire information display (Piolat et al., 1997), thus increasing users' cognitive load. The increased cognitive load should diminish the performance of tasks such as search accuracy and search speed.

For this study, I measured search accuracy by calculating the number of trials that participants correctly identify the correct result entry divided by the total trials. Search speed was recorded by using the timer in Qualtrics, and a self-report question was used to measure search task difficulty.

Based on the previous literature, I propose the following hypotheses addressing search accuracy, speed, and perceived difficulty:

Hypothesis 1a: Using a scrolling interface on search engine results pages (SERPs) will elicit lower search accuracy than a swiping interface.

Hypothesis 1b: Using a scrolling interface on SERPs will lead to slower search speed than a swiping interface.

Hypothesis 1c: Using a scrolling interface on SERPs will cause participants to rate the search task as more difficult than will a swiping interface.

2.3.1.2 Effects of Tactile Interactions on User Experiences. Besides users' performance of the search task, I'm also interested in users' subjective experiences such as time perception and satisfaction with the interface. Previous studies showed that in a retrospective paradigm (i.e., participants were not informed of the time perception question ahead of time), increased cognitive load lead to overestimate of time (i.e., increased subjective duration to objective duration ratio) (Block, Hancock & Zakay, 2010; Khan, Sharma & Dixit, 2006). Since scrolling may lead to higher cognitive load than swiping, I propose the following hypothesis concerning time perception:

Hypothesis 2: Scrolling on SERPs will elicit longer estimates of time than will swiping.

Previous studies showed inconsistent results linking user satisfaction with scrolling versus swiping interface. For instance, Choi, Kirshner and Wu (2016) found that participants were more engaged and satisfied with the swiping compared with scrolling interface of a mobile shopping app. However, Kim et al. (2016) test user satisfaction with scrolling and swiping interface when performing search tasks on a mobile phone and found no differences. Thus, I raise the following research question about user satisfaction:

Research Question 1: How does self-reported user satisfaction differ when using a scrolling versus a swiping interface?

2.3.2 Adding Movement Orientation to The Impact on Information Processing

This section builds on the discussion to this point by adding the influence of movement orientation, to discuss the interaction effect of tactile interaction and movement orientation on both primary and secondary content, and users' subjective experiences.

2.3.2.1. Effects of Tactile Interactions and Movement Orientation on Media Content. There are few studies looking at the effect of movement orientation (horizontal versus vertical) while controlling tactile interaction. For example, Ren, Kang, Ryu, and Han (2017) found that high Need-For-Touch participants prefer horizontal-swipe than vertical-swipe products, while low Need-For-Touch participants didn't show such difference. Braganza et al. (2009) found that more participants prefer vertical scroll than horizontal scroll (16 vs. 8). However, their reading performance did not differ because participants were allowed to check their answers.

There is a lack of evidence that movement orientation itself would influence task performance, so I don't expect to see the main effect of movement orientation. But I expect an interaction effect of tactile interaction and movement orientation. To be more specific, I predict that the effect of horizontal swiping versus vertical swiping will not be meaningful. However, horizontal scrolling and vertical scrolling might influence task performance differently. During vertical scrolling, the y coordinate of each piece of information is changing all the time. While in horizontal scrolling, the x coordinate of each piece of information changes only when the horizontal scroll enables the next page to push the previous page to the left but not entirely out of sight (referring to Figure 6). After the next page pushes the previous page out of sight in the horizontal scrolling condition, the x and y coordinates for each piece of information on the page are fixed. Thus, the cognitive load for processing location information is smaller in horizontal scrolling condition compared with vertical scrolling condition. Similar to the justification of Hypotheses 1a-c, a higher cognitive load may lead to worse task performance, thus vertical scrolling will lead to lower search accuracy than horizontal scrolling while performing a search task on SERPs. The predicted interaction effect is as follows:

Hypothesis 3a: There will be a significant interaction between tactile interaction and movement orientation on search accuracy such that there will be no difference in the swiping condition, but for the scrolling condition, horizontal scroll will generate higher search accuracy than vertical scroll.

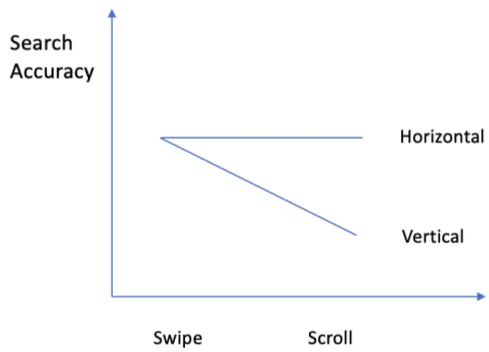


Figure 2. Illustration of proposed interaction effect on search accuracy (Hypothesis 3a)

Furthermore, in terms of search speed, vertical scrolling can only break the top (or bottom) line on the screen into parts that might impair reading, while horizontal scrolling can break all lines on the page into two parts (referring to Figure 6). Since we read line by line, when the horizontal scroll breaks all lines into parts, we cannot read efficiently (only reading the first couple words of each line does not make much sense). Users need to wait till the horizontal scroll fully reveals the next page and then read line by line. However, image processing and text processing are different. For images, people can scroll and recognize them faster, unlike text for which people need to finish scrolling in order to recognize whether a specific search term is the it's the target. Thus, I proposed a research question about search speed:

Research Question 2: How does search speed differ when using different combinations of tactile interaction (i.e., scrolling versus swiping) and movement orientation (i.e., horizontal versus vertical) interfaces?

In terms of the perceived difficulty of the search task, though horizontal scrolling has a lag time before the next page fully replaces the previous page to the left, the x and y coordinates are fixed when the screen only shows one-page content. However, the y coordinate for each information in the vertical scrolling is changing all the time, thus making it more difficult to represent the information in mind. I predict that users may find it more difficult to perform the search task in vertical scrolling condition than horizontal scrolling condition:

Hypothesis 3b: There will be a significant interaction between tactile interaction and movement orientation on perceived task difficulty such that there will be no difference in the swiping condition, but for the scrolling condition, vertical scrolling will lead to higher rating of task difficulty than horizontal scrolling.

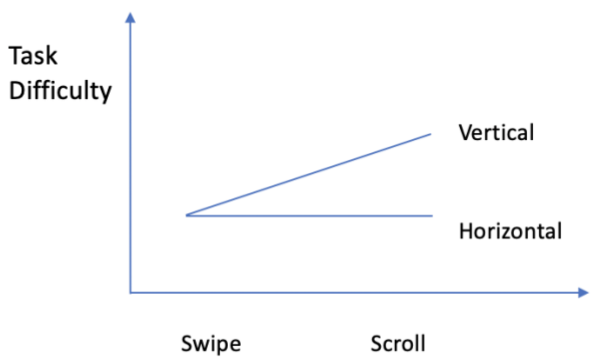


Figure 3. Illustration of proposed interaction effect on perceived task difficulty (Hypothesis 3c)

2.3.2.2 Effects of Tactile Interactions and Movement Orientation on User Experiences. Movement orientation itself cannot influence cognitive load. However, horizontal scrolling and vertical scrolling might differ in cognitive load (as the load to process location information is smaller in horizontal scroll compared with vertical scroll), thus influencing time perception. Therefore, I predict an interaction effect of tactile interaction and movement orientation:

Hypothesis 4: There will be significant interaction between tactile interaction and movement orientation on time perception such that there will be no difference in the swiping condition, but for the scrolling condition, vertical scrolling on SERPs will elicit longer estimates of time than will horizontal scrolling.

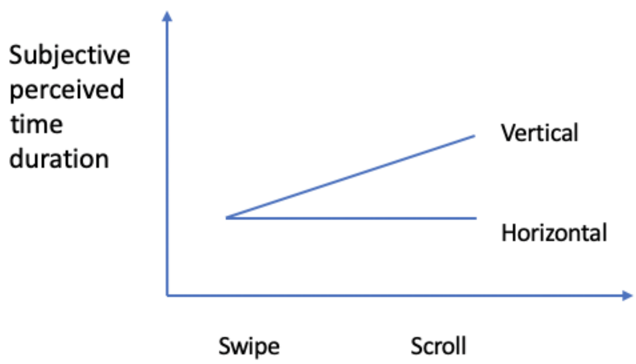


Figure 4. Illustration of proposed interaction effect on time perception (Hypothesis 4)

Because previous studies about tactile interaction and user satisfaction (mentioned in section 2.3.1.2) have been inconclusive, I proposed a research question about user satisfaction:

Research Question 3: How does self-reported user satisfaction differ when using different combinations of tactile interactions (i.e., scrolling versus swiping) and movement orientation (i.e., horizontal versus vertical) interfaces?

These hypotheses and research questions were tested in two studies in which people performed multiple search tasks on search engine result pages (SERPs). Their search accuracy, search speed, and perceived task difficulty were recorded during the task. Following the search tasks, their perception of task time, familiarity with using smartphones, and satisfaction about using the search interface were measured.

CHAPTER 3: STUDY ONE

3.1 Participants

276 participants were recruited through Amazon M-Turk to participate in the study for monetary compensation of $1.20. The compensation was calculated by using the federal minimum wage criteria and the expected time to complete the task, to ensure that it's no less than the federal minimum wage hourly rate.

3.2 Power Analysis

This sample size was based on a G*power analysis for the between-subject factor of a repeated measures ANOVA (F test). For the input parameters, effect size f was set as 0.2 (corresponding to a partial eta-squared value of 0.04), power was set as 0.8, the number of groups was set as 4, and the number of measurements was 8 to match the number of trials. This analysis yielded a recommended sample size of 160. To clarify, task performance measurements like search accuracy, search speed, and perceived task difficulty are repeated measures (8 trials). Self-reported measurements like time perception and user satisfaction were measured only once. The reason to use repeated measurements instead of non-repeated measurements to calculate sample size was due to the nature of the study. The study asked participants to perform search tasks using different interfaces, so I prioritized measuring task performances over self-reported satisfaction, thus leading to the use of repeated measures to calculate the sample size.

I ended up recruiting 276 participants from Amazon M-Turk because I found some participants failed to follow the instruction to perform the task or just entered random number for all the questions. I rejected those participants and opened the study to new participants to get enough valid data that matched the power analysis. This development is described in greater detail in forthcoming sections 3.7 and 3.8.

3.3 Design and Independent Variables

This experiment was a 2 (Tactile Interaction: scrolling versus swiping) × 2 (Movement Orientation: horizontal versus vertical) × 8 (Trials) mixed design. Tactile Interaction and Movement Orientation were between-subject variables while Trials was a within-subject variable. This design yielded four between-subjects conditions: horizontal swiping (each horizontal swipe revealed a new page that replace the old one), vertical swiping (each vertical swipe reveals a new page that replaces the old one), horizontal scrolling (each horizontal scroll reveals a certain amount of new information and push the same amount of old information out of sight), and vertical scrolling (each vertical scroll reveals a certain amount of new information and push the same amount of old information out of sight). In terms of technical implementation, scrolling mode tracks users' finger movements and allows the upcoming information to replace the current information corresponding to the speed of users' finger movement. In swiping mode, the replacement of one page of results with the next page lasts for 500 milliseconds from the time participants' fingers touch the screen, swipe through it, and leave the screen.

3.4 Task Stimuli

Each search task revealed twenty search results, separated in five pages (four items per page). The search tasks were based on previous research by Kim et al. (2016), only I used pictures instead of text results. An example can be found in Figure 5. The target search result could potentially appear in any of the twenty positions. More specifically, before participants used the interface to navigate through 20 products, they read a hypothetical scenario case: "Suppose you went to your friend's house and you would like to buy the exact same coffee maker that your friend had as the following picture. Please identify the exact coffee maker in the browsing list on the next page." One of the coffee makers (range from 1 to 20) would show together with the hypothetical scenario text, and any of the 20 products had the same probability to appear as the target.

Four conditions illustration:

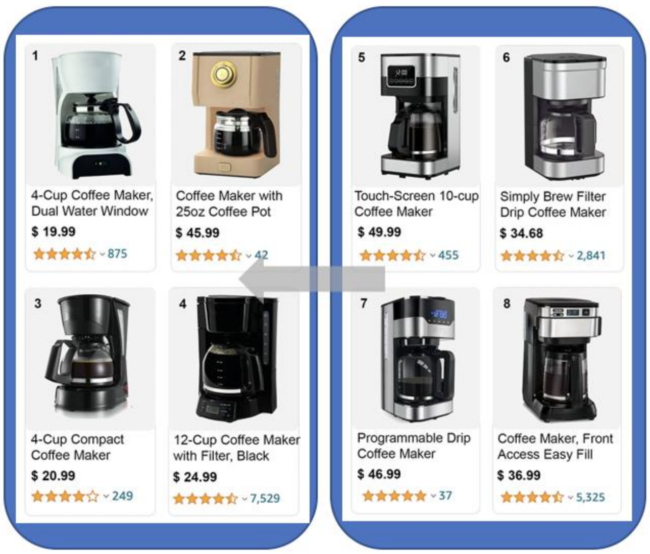


Figure 5. Horizontal swipe condition (one horizontal swipe from right to left makes the second page completely replace the first page) The grey arrow in all figures represents the direction of finger movements.

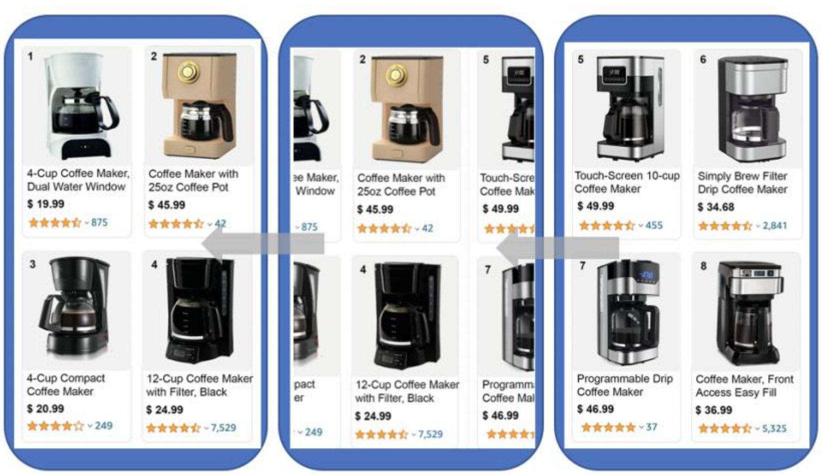


Figure 6. Horizontal scroll condition (the second page gradually pushes the first page to the left while scrolling)

Figure 7. Vertical swipe condition (one vertical swipe from bottom to top makes the second page completely replace the first page)

Figure 8. Vertical scroll condition (the second page gradually pushes the first page to the top while scrolling)

Figure 9. Four conditions for comparison

3.5 Control Variable

Power usage. Power usage is a measurement of users' familiarity with technologies. It served as a control variable because some of the dependent variables such as search speed and perceived task difficulty may not only influenced by interfaces, but also participants' familiarity with using technologies. Using power usage as a control variable can help to rule out this confounding variable.

When I conducted a literature review for power usage, I found Marathe et al. (2007) developed a power usage survey that has 24 items (some were adapted from the Technology Acceptance Model and others were developed based on literature) to measure individuals' technology usage in four aspects: motivation, expertise, efficacy, and behavior. However, the power usage survey developed by Marathe et al. (2007) is a general scale that measures an individual's familiarity with using technologies. Since this study was specific about navigating through different interfaces on smartphones, I used a more specific scale about measuring familiarity with using smartphones, which was developed by Kang and Shin (2016). It had three items which are specific about using smartphones: "I feel comfortable using various features on my smartphone", "I consider myself knowledgeable about smartphone use", "I know how to manage and control my personal information on my smartphone". All three items were on a 7-point scale anchoring between Strongly Disagree and Strongly Agree.

I first tested whether power usage has an individual effect on the outcome, then included it in the model to see if there were significant interaction effects with other variables. I didn't include it in the final model because the results showed that there were no individual effect nor interaction effects of power usage on outcome variables.

3.6 Dependent Variables

Search speed. Participants skimmed through 20 search results and identified the one that exactly matched the product they saw in the question. The time (in seconds) from the point at which the search result page appeared to the time when the participant specified a result entry was recorded as search speed automatically by using a timer which was invisible to participants in Qualtrics. Figure 10 showed an example of the task page.

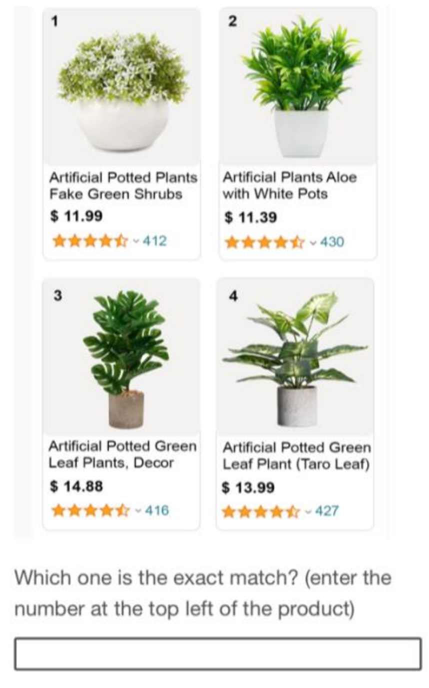


Figure 10. An example of the task page.

Search accuracy. The search accuracy was assigned the value of 1 when a participant identifies the correct/target result that exactly matched with the question; otherwise, it was assigned the value of 0. Since there were eight trials for each participant, the final search accuracy was the average search accuracy across all eight trials.

Perceived difficulty of the task. The purpose of asking the difficulty of the task in the main experiment is to test whether tactile interaction and movement orientation influence the perceived difficulty of the task. The question was "Please rate how difficult this task was for you by selecting one of the numbers below. A rating of 1 indicates that the participant felt the task was extremely easy whereas a rating of 7 indicates that the participant felt the task was extremely difficult" (Batho, Martinussen &Wiener, 2020).

Subjective perception of time (SPT). Subjective perception of time refers to users' estimation of the passage of time during an activity, compared against the actual time (Attfield et al., 2011). In research about attention, subjective perception of time is an indicator of cognitive aspects of engagement (Baldauf, Burgard & Wittmann, 2009). The more cognitive engaged someone is, the more likely that they would underestimate the pass of the time (Attfield et al., 2011). In previous research, subjective perception of time had been measured with categorical scale (Penney, Gibbon & Meck, 2008; Buetti & Lleras, 2012), interval scale (Sackett et al., 2010; Chinchanachokchai, Duff & Sar, 2015), and ratio scale (Grondin & Plourde, 2007; Luo et al., 2017). If using a categorical scale, participants were first taught to distinguished between a short and a long standard duration, then judge whether some given durations was closer to the short or the long standard one (Buetti & Lleras, 2012). If using an interval scale, participants were asked to use a Likert scale (0=very slowly, 10=very fast) to indicate how time seemed to progress for them during the task (Chinchanachokchai, Duff & Sar, 2015). While using a ratio scale, participants were asked to estimate in minutes and seconds, to the nearest 10s, of the perceived minimal and maximum duration of each task (Grondin & Plourde, 2007).

This study used "time passage" and "time duration" to measure time perception from previous literature (Xu & David, 2018). Time passage is a psychological perception of time along a bipolar continuum, which is tied to the limited capacity of attention (Lang, 2000). If a concurrent nontemporal task is demanding for attention, there would be less attentional resource for temporal processing, resulting in perception of speedy passage of time (Sucala et al., 2011).

Time duration is the subjective duration in seconds, minutes or hours, which is related to memory (Block & Reed, 1978). Time passage was measured by two interval questions asking participants to rate ‘‘Time appeared to go by very quickly when I was using the website’’ and ‘‘I lost track of time when I was using the website’’ on a 7-point scale (Novak, Hoffman, & Yung, 2000). Time duration was measured by a ratio-level question asking participants to estimate the actual duration in minutes and seconds of the search tasks.

User satisfaction. In usability studies, user satisfaction had been measured with a ten-item System Usability Scale (SUS), four-item Usability Metric for User Experience (UMUX), and two-item UMUX-LITE scale. Researchers in human-computer interaction have found these scales to be both reliable and strongly correlated to each other (Finstad, 2010; Lewis & Sauro, 2009, Lewis, Utesch & Maher, 2013, Borsci et al., 2015). I used the UMUS-LITE scale, which asks participant to rate “[This system’s] capabilities meet my requirements” and “[This system] is easy to use” from 1 (strongly disagree) to 7 (strongly agree) (Lewis, Utesch & Maher, 2013) to measure user satisfaction.

3.7 Pretest Results

Before the main experiment, I recruited 40 participants from M-Turk to do a pretest (10 for each condition). 7 participants from vertical swipe condition, 10 from vertical scroll condition, 9 from horizontal scrolling condition, and 10 from horizontal swiping condition finished the study; the other four didn’t finish the study. The average time they spent on the task was 7.2 minutes. I found 11 of them had a search accuracy lower than 50%. I then investigated these results. The following Table 2 shows the number of participants with disqualified data in each condition as well as the corresponding reason. All of them came from horizontal scroll/swipe conditions, and the majority of them were due to failing to navigate through all products as instructed.

Table 2

Number of Disqualified Participants in Each Condition in Pretest

	Didn't navigate through all products	Entered random number	Just performed badly	Total
Horizontal Scroll	5	2	0	7
Horizontal Swipe	3	1	0	4
Vertical Scroll	0	0	0	0
Vertical Swipe	0	0	0	0

I also had an open question at the end asking participants whether they had any difficulties when completing the study or anything they would like to mention about the study. I checked the response of those disqualified results and found they didn't enter anything specific (mostly just left blank or entered "None"). Since the data was collected online, some of the participants might have ignored the instructions and failed to scroll/swipe the product list horizontally to view all the products, thus resulting in such a low accuracy rate. I revised the instructions by turning the sentence that emphasizes the important points (i.e. there are 20 items and they can scroll/swipe horizontally/vertically to view all of them) bold and red. I also added a timer in Qualtrics that participants have to stay on the instruction page for at least 10 seconds before moving forward, to increase the chance that they read the instructions carefully. An example of the instructions can be found in Appendix A).

3.8 Experiment Procedure

For the main experiment, two hundred and seventy-six participants completed the study from Amazon M-Turk. Since there is no mechanism on Amazon M-Turk that prevents the same participant to sign up for multiple conditions, I launched four conditions sequentially and manually restricted those participants from previous sessions from signing up for multiple sessions. In order to rule out any possible order effect, I broke the data collection into two sets. For each set I randomized the order of four conditions and collected half of the data required for that particular condition. A detailed illustration of the data collection process on Amazon M-Turk can be found in Figure 11. I also restricted the participants' country to United States out of concern that participants from server farms in some countries generate low quality data (Moss & Litman, 2018).

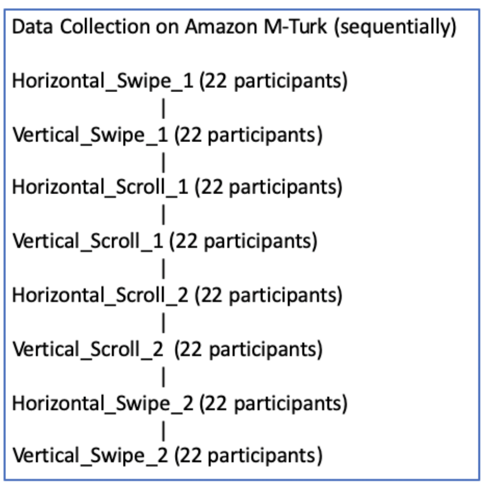


Figure 11. Data collection sequence of four conditions on Amazon M-Turk

Each participant signed up for one of four conditions: horizontal swiping, vertical swiping, horizontal scrolling, or vertical scrolling. For each condition, participants performed two practice trials and eight search tasks. Since most of the participants used laptops to sign up for M-Turk studies, I added the following sentence in the instructions: "In this task, you will be asked to conduct some online product searches using your smartphone with an operating system of either Android 7 or above, or IOS 13 or above." For those M-Turk workers who decide to participate in the study, they need to transmit the survey link from their laptops (if they used laptops to browse M-Turk studies in the first place) to their smartphones and open the survey link via smartphone. The link directed participants to a Qualtrics survey page. In Qualtrics, participants first read and e-signed a consent form, which explicitly said that they would need to use a smartphone with Android 7 or above, or IOS 13 or above (because older versions cannot experience the interface as expected). If they found out that their device didn't match the requirement, they were asked to quit the study and look for other studies on M-Turk. I also set a device identifier in Qualtrics after the consent form page. If any participants used laptops, tablets, or smartphones with older versions, they would reach the end of the survey right after the consent form with a message saying that the device they used was not qualified. In the consent form, I also specified that participants would only receive compensation if their search accuracy was greater than 50%. I set this criterion because the pretest results showed some participants just randomly entered a number for all questions. In order to prevent any bots or participants gaming the system to receive compensation without paying attention to the task, I asked them to perform at least half of the search task correctly to receive compensation.

After signing the online consent form, participants received the instruction about how to perform the tasks in the study. In total, they performed two practice trials and then eight search trials using their smartphones. The order of the eight target trials was randomized and the target item could appear in any position from 1 to 20 across five pages. After finishing each search task, participants received a short survey measuring their perceived task difficulty for that specific trial. Then they moved on to the next trial. After participants finished all the search tasks, they answered some questions about time perception, power usage and user satisfaction. Figure 12 is an illustration of experiment procedure.

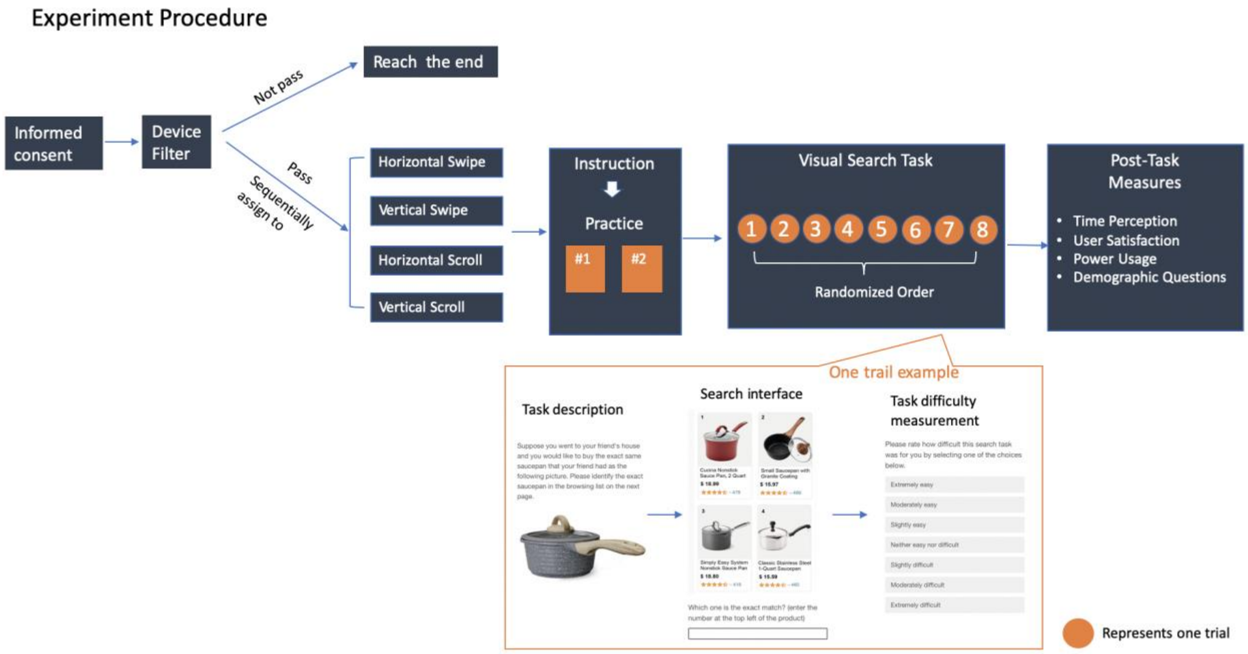


Figure 12. Experiment procedure

3.9 Data Preprocessing

I noticed that in total about 30.63% of M-Turk participants had a search accuracy lower than 50% because they either randomly entered a number for all trials (i.e. same number for all trials, number with decimal points, or numbers out of 1-20), didn't read the instruction carefully to scroll/swipe through the product lists to view all the items (i.e. only entered number from 1 to 4) or just perform badly (i.e. entered number between 1-20 but the accuracy was still lower than 0.5, which may due to paying attention to some of the tasks but failed to pay attention to the rest of the tasks). Table 3 below shows the number of the disqualify data in each condition. Since these participants didn't follow the instructions to perform the tasks, their data were unusable and had been removed from data analysis.

Table 3

Number of Disqualified Participants in Each Condition

	Didn't navigate through all products	Entered random number	Just perform badly	Total
Horizontal Scroll	44	15	3	62
Horizontal Swipe	8	7	1	16
Vertical Scroll	0	2	1	3
Vertical Swipe	0	2	0	2

Table 3 showed that far more participants failed to perform the task in horizontal conditions than vertical conditions, and relatively more in the horizontal scroll condition than vertical scroll condition, which indicated a systematic error. This might due to horizontal movement, particularly scrolling, being less familiar or intuitive than vertical movement in a visual search shopping scenario. Even if participants ignored the instructions in the vertical conditions, it stands to reason that they might have intuitively navigated the product list vertically and could perform the task correctly. But for participants in horizontal conditions, if they ignored the instructions and tried to scroll/swipe vertically and found it didn't work, they might mistakenly assume there were only four products in the product list to perform the task. If only looking at horizontal conditions, horizontal scroll had more invalid data points than horizontal swipe, which is likely also due to users' different level of familiarity of these two interactions. Horizontal swipe has existed for a long time since it's similar to flipping a page and had been used previously and extensively in eReader and weather apps. Horizontal scroll just came up in recent years and had been used in Google search results as a featured section along a continuous band or some news apps. Generally speaking, users might be less familiar with horizontal scroll and thus felt less intuitive to use it if they ignored the instructions, which cause the largest number of invalid data.

There was one invalid case for both time duration and time perception data (not the same participant). These cases were removed from analysis. Both the actual time spent on task and perceived time duration were in minutes. A new variable call “time duration” was created by using the estimated duration divided by the actual time they spent (Xu & David, 2018). This measure provides an intuitive ratio, such that any value greater than 1 represents overestimation of time duration and any value less than 1 represents underestimation. I used boxplots to detect outliers and used inter quarter range criteria to remove outliers for time duration. I reported the results after removing outliers in the following section. Time passage was calculated by the average of the two interval questions asking participants to rate ‘‘Time appeared to go by very quickly when I was using the website’’ and ‘‘I lost track of time when I was using the website’’ on a seven-point scale (Novak, Hoffman, & Yung, 2000). User satisfaction was calculated by aggregating two questions mentioned in the dependent variable section.

Besides measuring the total time participants spent on all the tasks, I also measured the time they spent on each trial. Since the distribution of time spent on each trial was highly positively skewed, I applied a log transformation of the time by using the formula log(1/reaction time) (Whelan, 2010) and ruled out the outliers based on the three standard deviation criterion. There were only 7 (out of 1496) trials outside the 3 standard deviation that had been removed. Then I ran the analysis on the transformed data.

.

3.10 Analyses and Results

After removing invalid data, one hundred and eighty-seven participants were included in the data analysis. I examined participants' search accuracy, search time, and perceived task difficulty using mixed-effects modeling. The continuous data were analyzed in liner mixed-effects models and a binomial dependent variable of search accuracy was analyzed in logistic mixed-effects models using the *lme4* package of R (Bates, Maechler, Bolker, & Walker, 2014; R Core Team, 2016). All models were fitted with a maximal random structure, including tactile interaction (swipe vs. scroll), movement orientation (vertical vs. horizontal), and interaction of tactile interaction and movement orientation as main effects, and subjects and trials (target product appear page) as random intercepts and slopes (Barr, Scheepers, & Tily, 2013). Means and standard deviations of the search accuracy, search time, and perceived task difficulty across conditions can be found in Table 4.

Table 4

Means and Standard Deviations of Task Performance Data

Tactile Interaction	Movement Orientation	Search Accuracy	Search Time	Perceived Difficulty
Scroll	Vertical	0.92 (0.27)	15.2 (8.16)	3.06 (2.37)
	Horizontal	0.94 (0.25)	14.0 (8.94)	2.84 (2.17)
Swipe	Vertical	0.92 (0.27)	15.3 (7.76)	2.86 (2.21)
	Horizontal	0.90 (0.30)	15.8 (10.00)	3.01 (2.22)

Note: Standard deviations in parentheses. Search time in seconds.

Hypothesis 1a predicted that using a scrolling interface on search engine results pages (SERPs) would elicit <u>lower search accuracy</u> than a swiping interface.

Logistic mixed-model analyses revealed that there are no significant effects of tactile interaction on search accuracy. The search accuracy in scrolling condition (M = 92.5% correct) is the same in swiping condition (M = 91.5% correct). This is probably indicative of a ceiling effect. Table 5 presents the model details and Figure 13 illustrates the accuracy across different conditions. The marginal R^2 is 0.01, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.46, which estimates the variance explained by all fixed effects and random effects in the model.

Table 5

Fixed-Effects of Logit Mixed-effects Model for Search Accuracy

	Estimate	SE	*z* value	*p* value
Intercept	3.69	0.50	7.34	< 0. 001
Tactile Interaction	-0.69	0.47	-1.46	0.14
Movement Orientation	-0.32	0.47	-0.67	0.49
Tactile Interaction x Movement Orientation	0.68	0.66	1.03	0.31

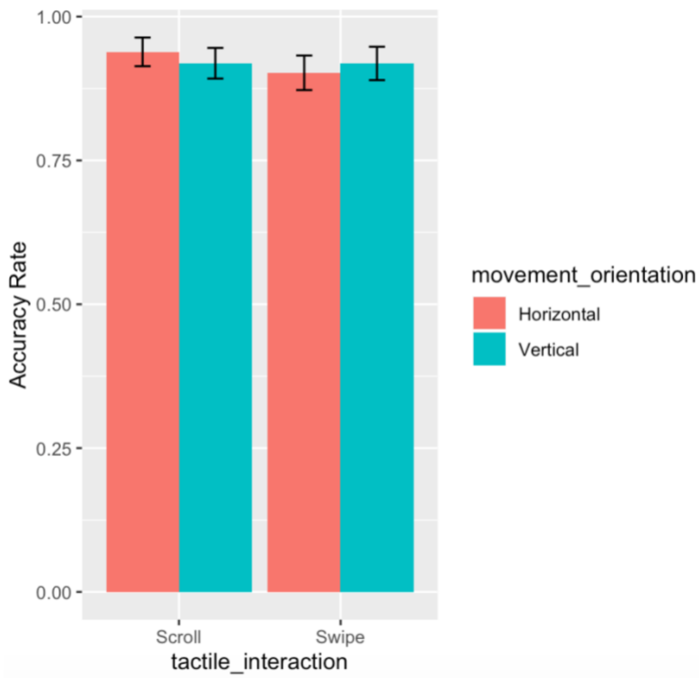


Figure 13. Search accuracy of tasks. Error bars represent 95% CI.

Hypothesis 1b predicted that using a scrolling interface on SERPs would lead to <u>slower search speed</u> than a swiping interface.

After applying a log transformation of the search time, linear mixed-effects model results revealed that the tactile interaction and movement orientation significantly influenced the search time. Participants spent less time searching when they use scrolling mode (M = 14.60) than when using swiping mode (M = 15.55). The results of search time show an opposite direction of Hypothesis 1b. Table 6 presents the fixed effects of the model results, and Figure 14 shows the inverse transform of log reaction time across conditions. The marginal R^2 is 0.01, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.41, which estimates the variance explained by all fixed effects and random effects in the model.

Table 6

Fixed-Effects of Linear Mixed-effects Model for Search Time

	Estimate	SE	*t* value	*p* value
Intercept	-2.49	0.09	-28.60	< 0.001
Tactile Interaction	**-0.12**	**0.06**	**-2.02**	**< 0.05**
Movement Orientation	**-0.12**	**0.06**	**-2.02**	**<0.05**
Tactile Interaction x Movement Orientation	0.11	0.09	1.23	0.22

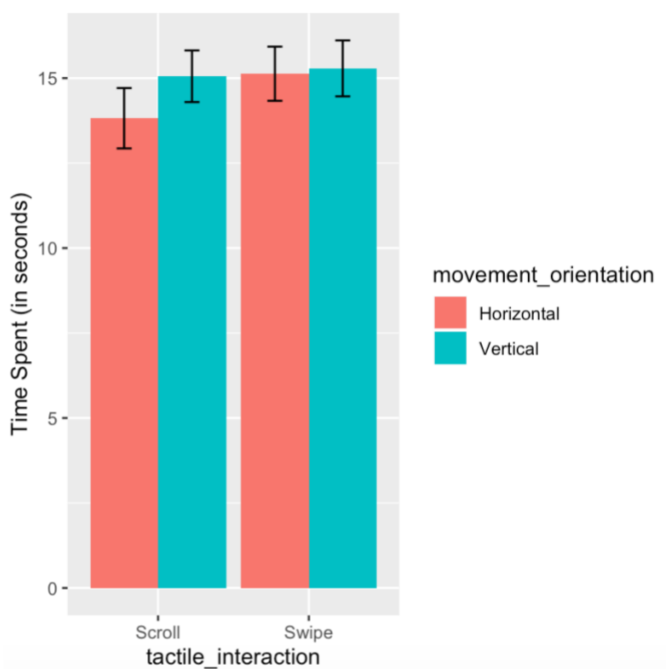


Figure 14. Search time of tasks after inverse log transform across conditions. Error bars represent 95% CI.

Since the target item was randomly showed on position 1-20 across five pages, there were some trials showed the target on the first page. For trials that showed the target on the first page, participants don't need to scroll/swipe to identify the correct match, thus the manipulation of tactile interaction and movement orientation didn't take place. Also, when the target was on the first page, participants would spend less time to identify the correct answer since they don't need to scroll/swipe. In order to further understand how this may have influenced the results, I decided to explore the dataset after deleting the trials that showed the target items on the first page (i.e. position 1 to 4) and re-run the analyses. Same as above, the results of search time after removing trials showed targets on the first page showed participants spent less time on task while using scrolling mode compared with swiping mode, which was opposite to hypothesis 1b. Table 7 presents the fixed effects of the model results and Figure 15 showed the inverse transform of log reaction time after deleting trials showed target on the first page. The marginal R^2 is 0.03, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.37, which estimates the variance explained by all fixed effects and random effects in the model.

Table 7

Fixed-Effects of Linear Mixed-Effects Model for Search Time After Removing Trials Showed Target On The First Page

	Estimate	SE	*t* value	*p* value
Intercept	-2.55	0.07	-38.51	< 0.001
Tactile Interaction	**-0.13**	**0.06**	**-2.30**	**< 0.05**
Movement Orientation	**-0.11**	**0.06**	**-2.00**	**<0.05**
Tactile Interaction x Movement Orientation	0.10	0.08	1.22	0.22

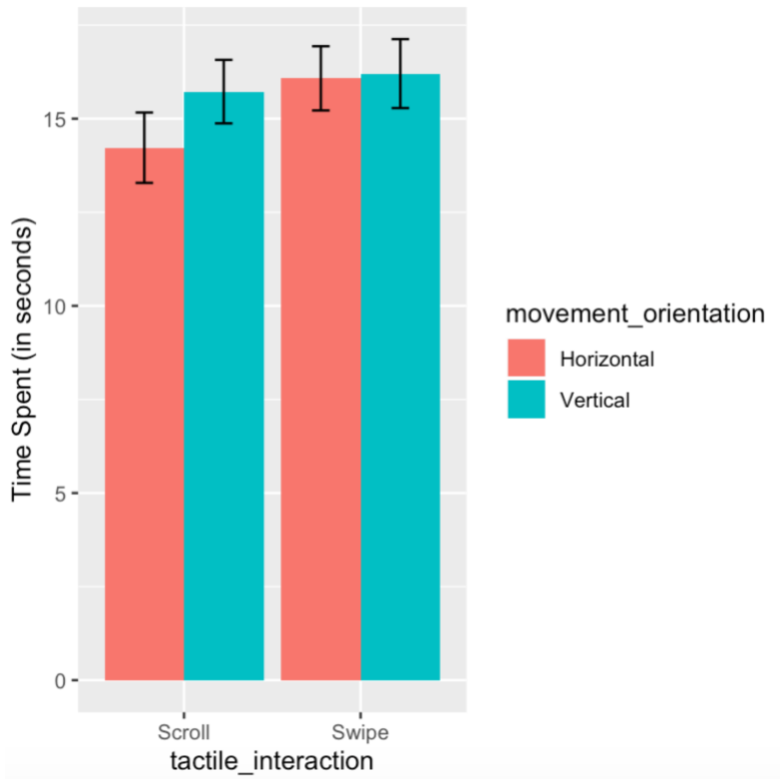


Figure 15. After removing trials that showed target on the first page, the search time of tasks after inverse log transform across conditions. Error bars represent 95% CI.

Hypothesis 1c predicted that using a scrolling interface on SERPs would cause <u>participants to rate the search task as more difficult</u> than will a swiping interface.

Analysis of the perceived task difficulty indicated both tactile interaction and movement orientation are not significantly affecting the perceived difficulty (Table 8). Subjects perceived them as "slightly easy" across all conditions. The marginal R^2 is 0.002, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.40, which estimates the variance explained by all fixed effects and random effects in the model.

Table 8

Fixed-Effects of Linear Mixed-effects Model for Perceived Task Difficulty

	Estimate	SE	*t* value	*p* value
Intercept	2.84	0.25	11.44	<0.001
Tactile Interaction	0.18	0.32	0.57	0.57
Movement Orientation	0.23	0.31	0.74	0.46
Tactile Interaction x Movement Orientation	-0.39	0.45	-0.87	0.39

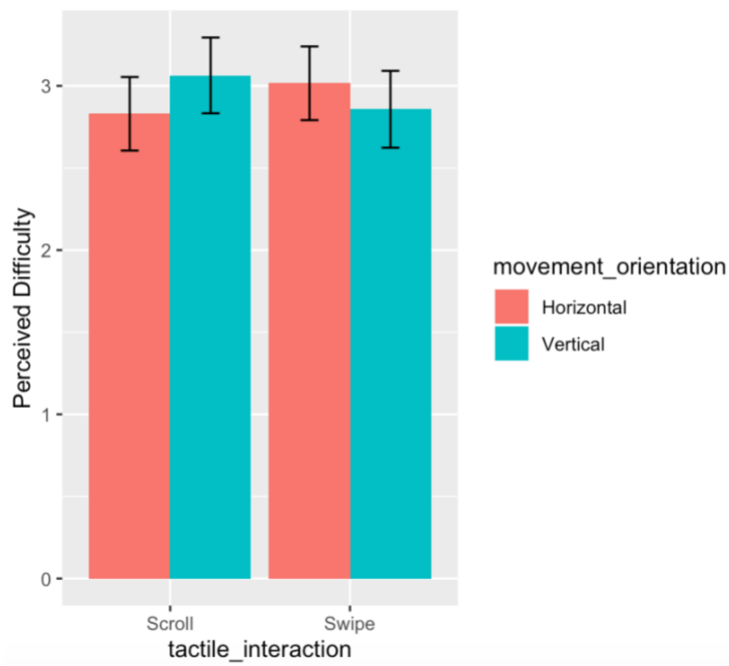


Figure 16. Perceived difficulty of tasks. Error bars represent 95% CI.

Hypothesis 2 predicted that scrolling on SERPs would prompt users to overestimate of time than swiping.

I measured both time duration and time passage. Time passage was the mean of two interval questions using a 7-point Likert scale measuring time perception from previous literature: ‘‘Time appeared to go by very quickly when I was using the website’’ and ‘‘I lost track of time when I was using the website’’(Novak, Hoffman, & Yung, 2000). Time duration is a ratio that used participants' estimation of time divided by the actual time. Table 9 presented the descriptive statistics of time passage and time duration across four conditions.

Table 9

Time Perception and User Satisfaction Data Descriptive Statistics

Tactile Interaction	Movement Orientation	Time Passage	Time Duration	User Satisfaction
Scroll	Vertical	4.82 (1.69)	0.86 (0.47)	5.69 (1.14)
	Horizontal	4.90 (1.33)	0.82 (0.47)	5.89 (0.90)
Swipe	Vertical	4.95 (1.49)	0.96 (0.42)	5.87 (0.96)
	Horizontal	4.88 (1.54)	0.72 (0.42)	5.73 (0.83)

Note: Standard deviations in parentheses.

A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA showed that tactile interaction didn't have significant effect on time passage, $F(1, 173) = .06$, $p > .05$. Similarly, the effect of tactile interaction on time duration was not significant, $F(1, 173) = 0.002$, $p > .05$. However, movement orientation showed a significant effect on time duration, $F(1, 173) = 4.05$, $p < .05$, partial - eta2 = 0.02, in which users in horizontal movement conditions (M = 0.77) tended to underestimate time duration more than users in vertical movement conditions (M = 0.91).

Hypothesis 3a predicted that there would be a significant interaction effect between tactile interaction and movement orientation on search accuracy such that there would be no difference in the swiping conditions, but for the scrolling conditions, horizontal scroll would generate higher search accuracy than vertical scroll.

Logistic mixed-model analyses revealed that there the interaction of tactile interaction and movement orientation on search accuracy is not significant (see Table 5).

Hypothesis 3b predicted that there would be a significant interaction effect between tactile interaction and movement orientation on search speed such that there would be no

difference in the swiping conditions, but for the scrolling conditions, vertical scroll would lead to faster search speed than horizontal scroll.

Linear mixed-effects model results revealed that the interaction of tactile interaction and movement orientation on search time is not significant (see Table 6).

Hypothesis 3c predicted that there would be a significant interaction effect between tactile interaction and movement orientation on perceived task difficulty such that there would be no difference in the swiping conditions, but for the scrolling conditions, vertical scroll would lead to higher rating of task difficulty than horizontal scroll.

ANOVA analyses revealed the interaction of tactile interaction and movement orientation on perceived task difficulty was not significant (see Table 8).

Hypothesis 4 predicted that there would be a significant interaction effect between tactile interaction and movement orientation on time perception such that there would be no difference in the swiping condition, but for the scrolling condition, vertical scroll would prompt users to overestimate time more than horizontal scroll.

Analysis of 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA showed that the interaction of tactile interaction and movement orientation on time passage and time duration were not significant, thus hypothesis 4 was not supported.

Research question 1 asks about how self-reported user satisfaction differ when using a scrolling versus a swiping interface. The descriptive statistics can be found in Table 9. A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA on user satisfaction showed that the effect of tactile interaction on user satisfaction was not significant, $F(1, 183) = .01$, $p > .05$.

Research question 2 asked how self-reported user satisfaction differ when using different combinations of tactile interactions (i.e., scrolling versus swiping) and movement orientation (i.e., horizontal versus vertical) interfaces. A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA on user satisfaction showed that the interaction effect of tactile interaction and movement orientation on user satisfaction were not significant, $F(1, 183) = 1.48$, $p > .05$.

3.11 Discussion

The goal of this study was to explore how tactile interaction (scrolling vs. swiping) and movement orientation (horizontal vs. vertical) influence users' task performance and user experience of different mobile interfaces. Results showed that users spent less time searching when using scrolling mode compared with swiping mode, while the search accuracy were the same (all above 90%). This result was contradicted with the hypothesis that using a scrolling interface on search engine result pages would lead to slower search speed than a swiping interface. Theoretically, positional information is processed when user consumes media content (Piolat et al., 1997, Lovelace & Southall, 1983, Baccino & Pynte, 1994), and scrolling mode would disrupt the spatial layout of media content while swiping mode wouldn't, which lead to higher cognitive load and thus slower search speed. The contradicted result probably due to participants were more used to the scrolling interfaces than the swiping interfaces in a shopping scenario, and this familiarity outweigh the different cognitive load placed by these two interfaces.

This result was also inconsistent with a previous study done by Kim et al. (2016), who found that participants spent longer to identify the correct search result when they used vertical scrolling compared with horizontal swiping. A possible reason is that my stimuli were pictures and their stimuli were textual sentences. In general, product pictures are less complex than textual search results, so the loads resulted from different interfaces might not differ extremely if using pictures than texts. It is possible that users' familiarity with the scrolling interface exceed the small amount of increased cognitive load and led to faster search speed.

I had an open question asking whether participants had any difficulties when completing the task or anything they would like to mention about the study at the end, and I received some useful comments that help me to refine the data collection process and instructions. One M-Turk participant entered "It might be better to provide a scan code to the survey, I ended up emailing the link to myself from the desktop because I don't usually do M-Turk stuff on mobile and logging into M-Turk from my phone is a pain". This comment inspired me to create an assignment survey to randomly show one of four QR codes that direct participants to a specific condition, which not only simplify the work for participants, but also allow me to collect data for four conditions at the same time.

I also received comments like "Some of the items didn't appear as options on the 4 numbers so I put zero as the answer", "There were a few items that did not show up in the choices afterwards. For example, there was a bee mat that it said to look for, but there was not a bee mat in the choice list afterwards. The bee mat was an obvious one, but it just wasn't there", "Most of the items did not match but there was a no match option". These comments indicated that some of the participants didn't read the instructions carefully so they didn't know they can scroll/swipe the list of products to view more. These comments inspired me to refine the instruction page by adding a gif picture showing the movement of the interface (e.g. horizontal scroll, vertical swipe, etc), in addition to the textual instructions. I believed the gif picture would be hard to miss even participants skip the textual instructions, which might help to lower the percentage of participants who fail to follow instructions.

CHAPTER 4: STUDY TWO

There was a high percentage of participants on Amazon M-Turk didn't follow the instructions of the tasks, especially for those who were assigned to horizontal scroll condition, which we can see from the Disqualified Participants in Each Condition (Table 3). Table 3 showed that there might be a systematic error in which more participants failed to follow the instruction in horizontal conditions than vertical conditions, and relatively more in horizontal scroll condition than vertical scroll condition. Based on the high percentage of invalid data collected from Amazon M-Turk, I decided to refine the instructions and collect another set of data from student sample to verify the trend.

4.1 Participants

172 undergraduate students at a large Midwestern university participated in the study for course credit.

4.2 Experiment Procedure

Study 2 used the same experiment design and same stimuli as Study 1 so the power analysis is the same, which revealed that 160 participants was needed. I adopted the same procedure with some revisions as Study 1. The differences were sampling from a different population, launching four conditions at the same time to ensure truly random assignment to conditions, and adding more details in the instructions to improve data quality.

Before I started to collect data from student sample, I created an assignment survey that showed one of four entries (i.e. QR codes students can use their smartphone to scan to enter a condition) to randomly assign participants. A detailed illustration of the data collection process for student sample in comparison with previous Amazon M-Turk sample can be found in Figure 17.

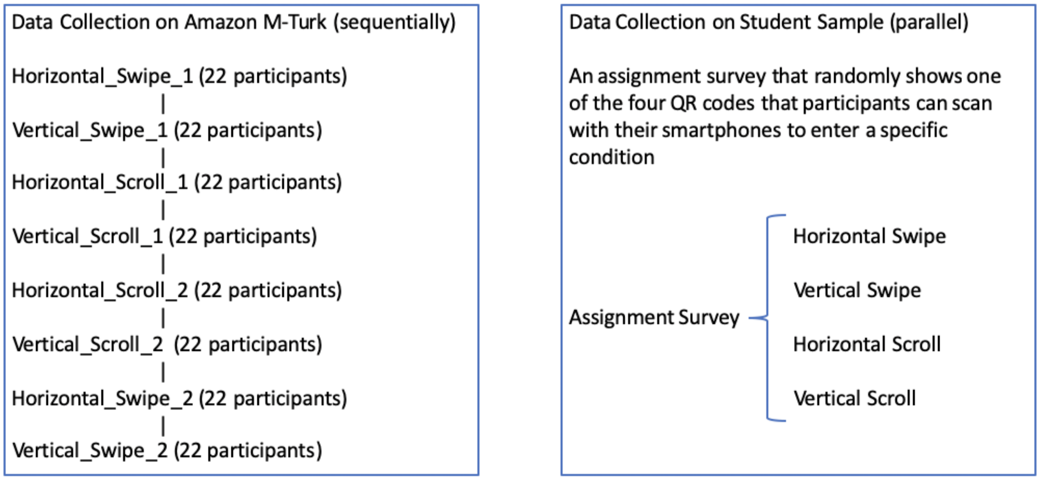


Figure 17. Data collection Process on Amazon M-Turk and Student Sample

I also refined the instruction by adding a gif picture showing the movement of the interface (e.g. horizontal scroll, vertical swipe, etc), to ensure they won't miss the important information about scrolling/swiping the interface to view the full list of products (An example of the instructions can be found in Appendix B). This help to lower the percentage of unusable data to 10.45%. The following Table 10 showed the number of disqualified data in each condition for student sample. In short, there were 172 students took part in the study, and 18 participants had a search accuracy lower than 0.5, which were mostly due to failure to follow the instruction to navigate through all the pages. This resulted in 154 usable cases for data analysis.

Table 10

Number of Disqualified Participants in Each Condition

	Didn't navigate through all products	Entered random number	Just perform badly	Total
Horizontal Scroll	11	2	0	13
Horizontal Swipe	1	0	1	2
Vertical Scroll	1	0	1	2
Vertical Swipe	0	0	1	1

Table 10 showed a similar trend with Table 3, indicating that more participants in horizontal scroll condition failed to follow the instructions to navigate through all products, compared with other three conditions.

4.3 Data Preprocessing

For the time perception data, there were seven participants that didn't enter a value of minutes and seconds, so I removed them from analysis for time perception. Both the actual time spent and perceived time duration were in minutes. A new variable call "time duration" was created by using the estimated duration divided by the actual time they spent (Xu & David, 2018). I used boxplots to detect outliers and used inter quarter range criteria to remove outliers for time duration. I ran analysis on the dataset before and after removing outliers and they told the same story, and I reported the results after removing outliers in the following section. On the other hand, time passage was calculated by the average of the two interval questions asking participants to rate ''Time appeared to go by very quickly when I was using the website'' and ''I lost track of time when I was using the website'' on a seven-point scale. Last but not least, user satisfaction was calculated by aggregating two questions mentioned in the dependent variable section.

Besides measuring the total time participants spent on all the tasks, I also measured the time they spent on each trial. Since the distribution of time spent on each trial was highly right skewed, I applied a log transformation of the time by using the formula log(1/reaction time) (Whelan, 2010) and rule out the outliers based on the three standard deviation criteria. There was only 6 (out of 1232) trials outside the 3 standard deviation that had been removed. Then I ran the analysis on the transformed data.

4.4 Analyses and Results

After removing invalid data, one hundred and fifty-four participants were included in the data analysis. I examined participants' search accuracy, search time, and perceived task difficulty using mixed-effects modeling. The continuous data were analyzed in liner mixed-effects models and a binomial dependent variable of search accuracy was analyzed in logistic mixed-effects models using the *lme4* package of R (Bates, Maechler, Bolker, & Walker, 2014; R Core Team, 2016). All models were fitted with a maximal random structure, including tactile interaction (swipe vs. scroll), movement orientation (vertical vs. horizontal), and interaction of tactile interaction and movement orientation as main effects, and subjects and trials (target product appear page) as random intercepts and slopes (Barr, Scheepers, & Tily, 2013). Means and standard deviations of the search accuracy, search time, and perceived task difficulty are found in Table 11.

Table 11

Means and Standard Deviations of Task Performance Data

Tactile Interaction	Movement Orientation	Search Accuracy	Search Time	Perceived Difficulty
Scroll	Vertical	0.94 (0.23)	11.30 (5.69)	2.09 (1.89)
	Horizontal	0.97 (0.16)	9.99 (4.32)	2.11 (2.03)
Swipe	Vertical	0.92 (0.27)	10.40 (5.69)	1.84 (1.62)
	Horizontal	0.92 (0.28)	10.50 (5.21)	2.52 (2.24)

Note: Standard deviations in parentheses. Search time in seconds.

Hypothesis 1a predicted that scrolling on search engine results pages (SERPs) would elicit <u>lower search accuracy</u> than swiping presentation.

Logistic mixed-model analyses revealed that there is a significant effect of tactile interaction on search accuracy. The search accuracy in scrolling condition (M = 95.5% correct) is higher than in swiping condition (M = 92% correct), which shows an opposite pattern of hypothesis 1a. Table 12 presents the model details and Figure 18 illustrates the accuracy across different conditions. The marginal R^2 is 0.05, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.36, which estimates the variance explained by all fixed effects and random effects in the model.

Table 12

Fixed-Effects of Logit Mixed-effects Model for Search Accuracy

	Estimate	SE	*z* value	*p* value
Intercept	4.36	0.55	7.87	< 0. 001
Tactile Interaction	**-1.27**	**0.56**	**-2.25**	**<0.05**
Movement Orientation	-0.82	0.56	-1.47	0.14
Tactile Interaction x Movement Orientation	0.83	0.72	1.15	0.25

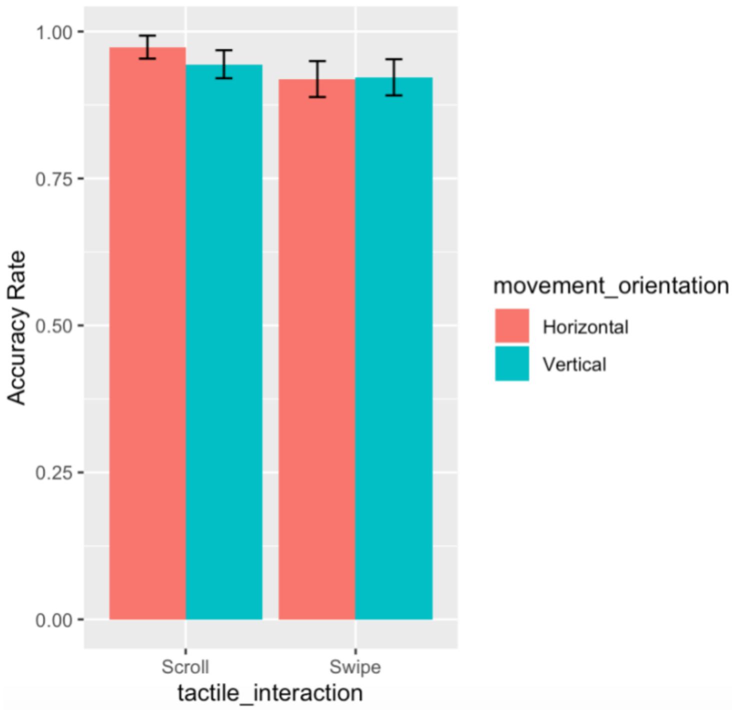


Figure 18. Search accuracy of tasks. Error bars represent 95% CI.

Hypothesis 1b predicted that scrolling on SERPs would lead to <u>slower search speed</u> than swiping presentation.

Linear mixed-effects model results revealed that tactile interaction does not have significant effect on search time. Participants spent similar time searching when they use scrolling mode (*M* = 10.65) or when using swiping mode (*M* = 10.45). Table 13 presents the model details and Figure 19 illustrates the search time across different conditions. The marginal R^2 is 0.003, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.43, which estimates the variance explained by all fixed effects and random effects in the model.

Table 13

Fixed-Effects of Linear Mixed-effects Model for Search Time

	Estimate	SE	*t* value	*p* value
Intercept	-2.24	0.10	-22.19	<0.001
Tactile Interaction	0.02	0.06	0.27	0.79
Movement Orientation	-0.05	0.06	-0.85	0.40
Tactile Interaction x Movement Orientation	0.04	0.09	0.47	0.64

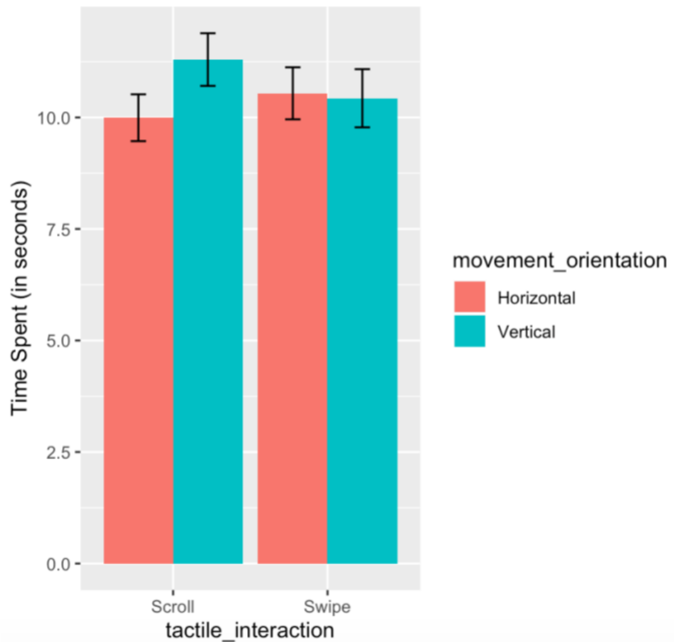


Figure 19. Search time of tasks. Error bars represent 95% CI.

Since the target item was randomly showed on position 1-20 across five pages, there were some trials showed the target on the first page. For trials that showed the target on the first page, participants don't need to scroll/swipe to identify the correct match, thus the manipulation of tactile interaction and movement orientation didn't take place. I decided to explore the dataset after deleting the trials that showed the target item on the first page (i.e. position 1 to 4) and re-run the analysis. Table 14 presents the fixed effects of the model results and Figure 20 showed the inverse transform of log reaction time across four conditions. The marginal R^2 is 0.004, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.31, which estimates the variance explained by all fixed effects and random effects in the model.

Table 14

Fixed-Effects of Linear Mixed-effects Model for Search Time After Removing Trials Showed Target on the First Page

	Estimate	SE	*t* value	*p* value
Intercept	-2.30	0.06	-40.71	<0.001
Tactile Interaction	-0.04	0.06	-0.71	0.48
Movement Orientation	-0.07	0.06	-1.33	0.18
Tactile Interaction x Movement Orientation	0.08	0.08	1.05	0.30

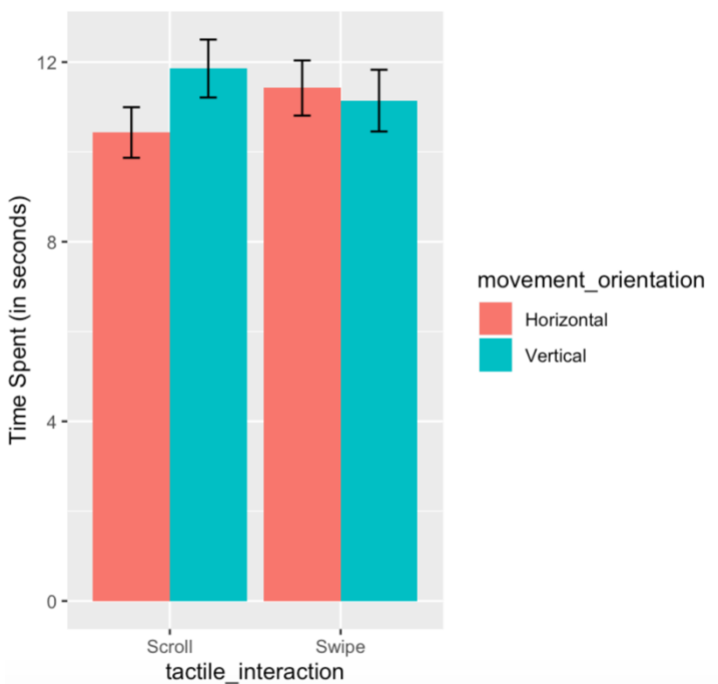


Figure 20. After removing trials that showed target on the first page, the search time of tasks after inverse log transform across conditions. Error bars represent 95% CI.

Hypothesis 1c predicted that scrolling on SERPs would <u>prompt participants to rate the search task as more difficult</u> than swiping presentation.

Linear mixed-effects model results revealed that tactile interaction did not have a significant effect on perceived task difficulty. Participants perceived the search tasks as similar difficulty level when they used scrolling mode (M = 2.10) or when using swiping mode (M = 2.18), indicating a likely floor effect. Participants perceived all the tasks as “moderately easy”. Table 15 presents the model details and Figure 21 illustrates the search time across different conditions. The marginal R^2 is 0.01, which is the variance explained by all fixed effects in the model, the conditional R^2 is 0.33, which estimates the variance explained by all fixed effects and random effects in the model.

Table 15

Fixed-Effects of Linear Mixed-effects Model for Perceived Task Difficulty

	Estimate	SE	*t* value	*p* value
Intercept	2.11	0.24	8.64	<0.001
Tactile Interaction	0.39	0.28	1.40	0.17
Movement Orientation	-0.02	0.27	-0.08	0.93
Tactile Interaction x Movement Orientation	-0.63	0.38	-1.65	0.10

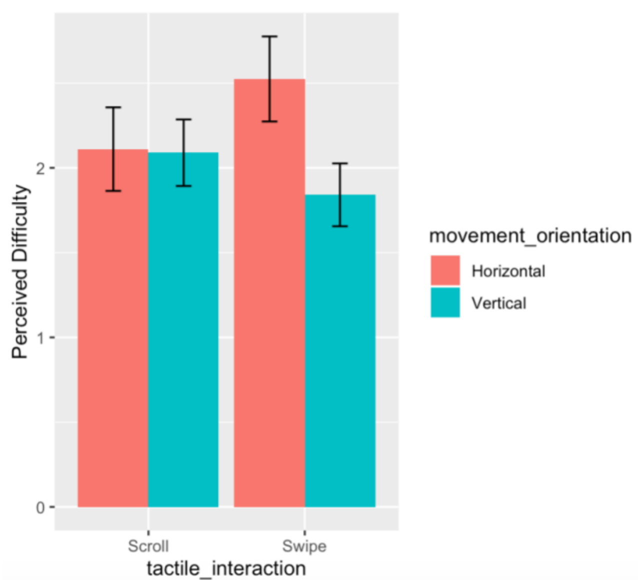


Figure 21. Perceived difficulty of tasks. Error bars represent 95% CI.

Hypothesis 2 predicted that scrolling on SERPs would prompt users to overestimate of time than swiping.

I measured both time duration and time passage. Time passage was the mean of two interval questions using a 7-point Likert scale measuring time perception from previous literature: ‘‘Time appeared to go by very quickly when I was using the website’’ and ‘‘I lost track of time when I was using the website’’(Novak, Hoffman, & Yung, 2000). Time duration is a ratio that used participants’ estimation of time divided by the actual time.

Table 16

Means and Standard Deviations of Time Perception and User Satisfaction Data

Tactile Interaction	Movement Orientation	Time Passage	Time Duration	User Satisfaction
Scroll	Vertical	4.59 (1.42)	0.68 (0.35)	5.49 (1.39)
	Horizontal	4.95 (1.41)	0.62 (0.28)	5.61 (1.19)
Swipe	Vertical	4.55 (1.40)	0.66 (0.30)	5.82 (0.94)
	Horizontal	4.49 (1.19)	0.64 (0.37)	5.54 (1.10)

Note: Standard deviations in parentheses.

A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA showed that tactile interaction didn’t have significant effect on time passage, $F(1, 138) = 1.15$, $p > .05$. Similarly, the effect of tactile interaction on time duration was not significant, $F(1, 138) = 0.000$, $p > .05$. Hypothesis 2 was not supported since tactile interaction doesn’t have significant effect on time perception.

Interestingly, by looking at the time duration data of Study 2 and Study 1, I found that student sample (from Study 2) tended to underestimate the task time more than M-Turk workers (from Study 1). Table 16 showed that on average student sample entered their perceived time as 62%-68% of the actual task time, while Table 9 showed that M-Turk sample entered their perceived time as 72%-96% of the actual task time.

Hypothesis 3a predicted that there would be a significant interaction effect between tactile interaction and movement orientation on search accuracy such that there would be no difference in the swiping conditions, but for the scrolling conditions, horizontal scroll would generate higher search accuracy than vertical scroll.

Logistic mixed-model analyses revealed that there the interaction of tactile interaction and movement orientation on search accuracy is not significant (see Table 12).

Hypothesis 3b predicted that there would be a significant interaction effect between tactile interaction and movement orientation on perceived task difficulty such that there would be no difference in the swiping conditions, but for the scrolling conditions, vertical scroll would lead to higher rating of task difficulty than horizontal scroll.

ANOVA analyses revealed the interaction of tactile interaction and movement orientation on perceived task difficulty was not significant (see Table 15).

Hypothesis 4 predicted that there would be a significant interaction effect between tactile interaction and movement orientation on time perception such that there would be no difference in the swiping condition, but for the scrolling condition, vertical scroll would prompt users to overestimate time more than horizontal scroll.

Analysis of 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA showed that the interaction of tactile interaction and movement orientation on both time passage and time duration were not significant, thus hypothesis 4 was not supported.

Research question 1 asked about how self-reported user satisfaction differ when using a scrolling versus a swiping interface. The descriptive statistics can be found in Table 16. A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA on user satisfaction showed that the effect of tactile interaction on user satisfaction were not significant, $F(1, 150) = .59$, $p > .05$.

Research question 2 asked about how search speed differ when using different combinations of tactile interactions (i.e., scrolling versus swiping) and movement orientation (i.e., horizontal versus vertical) interfaces. Linear mixed-effects model results revealed that the interaction of tactile interaction and movement orientation on search time was not significant (see Table 13).

Research question 3 asked about how self-reported user satisfaction differ when using different combinations of tactile interactions (i.e., scrolling versus swiping) and movement orientation (i.e., horizontal versus vertical) interfaces. A 2 (tactile interaction) x 2 (movement orientation) two-way ANOVA on user satisfaction showed that the interaction effect of tactile interaction and movement orientation on user satisfaction were not significant, $F(1, 150) = 1.11$, $p > .05$.

4.5 Discussion

The results from Study 2 showed that participants spent the same amount of time on search tasks across different interfaces but had higher search accuracy when using scrolling interface compared with swiping interface. This result was contradicted with the hypothesis that using a scrolling interface on search engine result pages would lead to worse search accuracy than a swiping interface. It also contradicted with a similar study done by Kim et al. (2016), who found that horizontal swiping resulted in better search accuracy than vertical scrolling when the correct answer located beyond the first page. A possible reason is that my stimuli were pictures and their stimuli were textual sentences. In general, product pictures are less complex than textual search results, so the loads resulted from scrolling vs. swiping interfaces might not differ extremely if using pictures than texts. It is possible that users' familiarity with the scrolling interface exceed the small amount of increased cognitive load and led to better search accuracy.

In addition, Study 2 and Study 1 showed different result patterns but they told the same general story: users perform search tasks better when using scrolling interfaces compared with swiping interfaces. More specifically, Study 1 showed that users spent less time searching when using scrolling mode compared with swiping mode, while the search accuracy were the same. Study 2 showed that users had higher search accuracy when using scrolling mode compared with swiping mode, while the search time were the same. The different emphasis (one for accuracy the other for search time) for these two samples might due to slightly different instructions about receiving compensations. In order to prevent M-Turk workers gaming the system to received money compensation without doing the tasks, I specified that they need to score an accuracy no less than 50% to receive compensation. Results from M-Turk sample showed the same high accuracy across conditions, but the time spent on different interfaces differ. However, for student sample, I didn't set a criterion for them to receive credits. Basically, they can receive credits if they participate, regardless of task performances. So, results from student sample showed they spent the same amount of time on tasks across different conditions, but their search accuracy differ. Table 17 lists the differences between Study 1 and Study 2.

In addition, by comparing time perception data between Study 2 and Study 1, I found that the student sample (from Study 2) tended to underestimate the task time more than M-Turk workers (from Study 1). Table 16 showed that on average student sample entered their perceived time as 62%-68% of the actual task time, while Table 9 showed that M-Turk sample entered their perceived time as 72%-96% of the actual task time. One possible explanation for this is that college students might perceive these search tasks as easier, relative to a more general population. By looking at the task performance Table 4 (from M-Turk sample) and Table 11 (from student sample), students completed the tasks faster than M-Turk sample (average 10-11 seconds per task for students vs. 14-16 seconds per task for M-Turk workers) while maintaining the same level of accuracy rate. Also, students perceived the task as easier than M-Turk workers (average 2 vs 3 on a 7-point scale). Studies found that in a retrospective paradigm (i.e., participants were not informed of the time perception question ahead of time), increased cognitive load lead to overestimate of time (i.e., increased subjective duration to objective duration ratio) (Block, Hancock & Zakay, 2010; Khan, Sharma & Dixit, 2006). Self-report task difficulty, task completion time, performance accuracy can be indicators of cognitive load (Chen, Epps & Chen, 2011). This implies that perhaps the tasks were easier for students, leading to less perceived cognitive load and greater underestimation of time, relative to the M-Turk sample.

Table 17

Differences between Study 1 and Study 2

	Study 1	Study 2
Sample	Amazon M-Turk Sample, 276 participants were recruited, resulted in 187 valid data	Student Sample, 172 participants were recruited, resulted in 154 valid data
Data Collection Process	Sequentially for four conditions while taking care of order effect	In parallel for four conditions
Instructions	Emphasized to score no less than 50% in search accuracy to receive compensation	Added a gif picture to illustrate how to swipe/scroll through the interface beyond textual instructions
Results	Same accuracy across four conditions. Faster search speed in scrolling conditions than swiping conditions	Same search speed across four conditions. Higher search accuracy in scrolling conditions than swiping conditions

CHAPTER 5: DISCUSSION

This study tested whether tactile interaction and movement orientation influence users' search task performance and user experience on mobile devices. This section provides some final interpretations of the study and suggestions for the human computer interaction (HCI) community.

5.1 Interpretation of the Data

Results from two studies indicated that users perform better while using a scrolling interface compared to a swiping interface: either higher search accuracy while the search time was the same, or faster search speed while the search accuracy was the same. This may be because users' familiarity of using scrolling interface in a shopping scenario outweigh the higher cognitive load placed by the scrolling interface compared with swiping ones. One previous study compared users' search performance while using a pagination (similar to swiping) or scrolling interface on a mobile phone, and found participants had better search speed and accuracy when using pagination (Kim et al., 2016).

This study adopted similar experiment design but used product pictures instead of textual search results and found opposite patterns, which indicated that the type of information (picture vs. text) also matter. The load time for text and images on the speed of judgment is different. For pictures, people can scroll and recognize it faster, but for text, people have to finish the scroll and decide whether it's the target. It's possible that for textual information, users search faster and more accurate while using swiping interfaces comparing with scrolling interfaces, but for pictorial information, users search faster and more accurate while using scrolling interfaces compared with swiping ones. This is also in line with the theory that cognitive load has three specific types: intrinsic (which is decided by the content), extraneous (which is everything else related to the information acquiring experience), and germane load (which is determined by individuals' effortful learning) (Schnotz & Kürschner, 2007; DeLeeuw & Mayer, 2008; De Jong, 2010; Kalyuga, 2011). Different tactile interactions might cause different extraneous load, different type of stimuli (picture vs. text) might cause different intrinsic load, different samples might have different germane load.

Using textual stimuli with different tactile interactions may vary the cognitive load between scrolling and swiping as much as possible, while using pictorial stimuli might not vary the load that much across conditions. This could be supported by participants' perceived task difficulty data, since self-report task difficulty is one of the measurements of cognitive load (Chen, Epps & Chen, 2011). In both Study 1 and Study 2, participants perceived task difficulty did not differ across conditions, which indicated that the cognitive load might not differ that much across different tactile interactions and movement orientation conditions for the same sample if using pictures as stimuli.

In addition, the student sample's perceived task difficulty data was a bit lower than M-Turk sample (2 vs. 3 on a 7-point scale), which indicated that the germane load was lower in student sample than M-Turk workers sample. Since they performed the same task with the same stimuli, which means the intrinsic load and extraneous were the same, the overall cognitive load of the task to student sample might be lower compared with the M-Turk worker sample. This could help to explain why student sample tended to underestimate the task time more than M-Turk workers (estimate time to actual time ratio was 62%-68% for students and 72%-96% for M-Turk workers). Previous literature showed that increased cognitive load leads to overestimate of time (i.e., increased subjective duration to objective duration ratio) when participants were not informed of the time perception question ahead of time (Block, Hancock & Zakay, 2010; Khan, Sharma & Dixit, 2006). In other words, when the cognitive load of the same task in student sample was lower compared with M-Turk sample, they would underestimate the time more, which was reflected in the time duration data.

There were no significant differences in ratings about satisfaction of interfaces across four conditions. This result was inconsistent with previous literature that showed swiping leads to higher cognitive absorption and playfulness than scrolling in a shopping scenario (Choi, Kirshner & Wu, 2016). One possible reason is that this study was a goal-oriented search and shop task, in which participant had to find a specific product among a list; and Choi, Kirshner and Wu's study was a more exploratory shopping task in which participant could add any item they like into a shopping cart. Participants may not spend much time experiencing the interfaces in a more goal-oriented search task compared with an exploratory task thus their satisfaction rating might not differ much across different interfaces.

However, from the invalid data that showed the number of participants who failed to navigate through the list of all products, horizontal scroll generated a large amount of invalid data compared with other three conditions (44 for horizontal scroll vs 8 for other three conditions combined in Study 1, 11 for horizontal scroll vs 2 for other three conditions combined in Study 2). I also got two comments from participants in the horizontal scroll condition saying, "I thought having to scroll sideways was a bit cumbersome", and "I prefer scrolling down versus sideways, but it was not difficult to follow". This might suggest that horizontal scroll is the least intuitive among the four conditions in a shopping scenario, and app designers should probably consider avoiding only using horizontal scroll while designing visual search or shopping interfaces.

I also ran some descriptive analyses of the invalid data for both the M-Turk sample and student sample. It showed that participants whose data were invalid spent less time, perceived the task to be more difficult, and tended to underestimate time less (reflected by time duration, which is calculated by estimate time/ actual time) compared with valid data. Detailed descriptive statistics can be found in Table 18.

Table 18

Means and Standard Deviations of Task Performance, Time Perception and User Satisfaction Data for Invalid Data

Sample	Task Performance Data		Time Perception and User Satisfaction Data		
	Search Time	Perceived Difficulty	Time Passage	Time Duration	User Satisfaction
M-Turk Invalid	10.99 (8.37)	3.36 (2.36)	4.96 (1.48)	0.96 (0.63)	5.46 (1.32)
M-Turk Valid	15.08 (8.72)	2.94 (2.24)	4.89 (1.51)	0.84 (0.45)	5.80 (0.96)
Student Invalid	7.49 (3.26)	4.51 (2.45)	4.17 (1.04)	0.79 (0.58)	4.78 (1.18)
Student Valid	10.55 (5.23)	2.14 (1.95)	4.65 (1.36)	0.65 (0.33)	5.62 (1.16)

Since most of these invalid data were due to failure to navigate through all five pages to identify the target product, it’s reasonable that they spent less time on each search task (because they didn’t scroll/swipe through all pages), perceived the task to be more difficult (because they didn’t find the target), and estimate the time as more accurate to the actual time. However, the satisfaction with the interface did not differ much between the invalid and valid data for M-Turk sample (5.46 for invalid and 5.74 for valid) but appeared to be lower for invalid data for student sample (4.78 for invalid and 5.62 for valid). However, it’s inappropriate to claim any of these findings is significant since the invalid data for student sample was 18 and the invalid data for M-Turk sample was 83.

5.2 Limitations

These findings cannot be interpreted without considering several possible limitations. One of the limitations of this study is that the familiarity with these four different interfaces (horizontal swiping, horizontal scrolling, vertical swiping, vertical scrolling) was not measured or controlled. Although all these four types of interactions had been embedded in mobile webs or apps (e.g. weather apps using horizontal swiping, Tiktok using vertical swiping, Google search featured results sections using horizontal scrolling, and normal webpage using vertical scrolling), users' familiarity with them might still differ. Future research should measure not only users' familiarity with using smartphones, but also their familiarity with using specific interfaces so that these variables can be controlled.

Another limitation of the study is that the way participants specify the target is not that intuitive. Since there were data transfer privacy issues between Qualtrics and its embedded external websites, participants cannot just tap the product and go next. Instead, they need to enter the number of the correct product in Qualtrics after navigating through the embedded website, which is somewhat artificial and not that intuitive. In addition, the time participant spent on each task was measured from moment that the product page appeared to them until they entered the target number and click the "next" button, which would be longer than just tapping a product and go next. Also, the scrolling interfaces track users' finger speed and allow the new information to replace the old information corresponding to users' finger speed, while swiping interfaces have a constant speed of 500 milliseconds to replace the old page with a new one. All of these make the search speed less sensitive to the differences caused by different interfaces. Future research should use a more advanced app that includes a seamless timing mechanism and allow finger speed tracking across all interfaces.

Another limitation of the study is that it’s difficult to verify the effect of spatial encoding in this study. Although the spatial encoding load differs between swiping and scrolling interfaces, it also varies by task. For search tasks to identify a target image, users first saw the target, then browsed a product list to identify the target. They might have stopped searching right after they found the target in the product lists, thus making spatial codes less relevant. Spatial codes might only play a role when users were not sure about which item is the target and scrolled or swiped back and forth to compared different items. However, it’s difficult to verify a user’s actual behavior without tracking or recording it. Future studies can address this issue by asking permission from users to track/record real-time behavior data for further analysis.

The generalizability of the study is limited. This study only investigated a very small realm of interactions: scrolling versus swiping when users perform visual search tasks. In addition, to make four conditions comparable, I created an artificial horizontal scroll in which the whole page gradually pushes the old information from right to left; while in real life horizontal scroll is applied to only a band of information on the page, as mentioned in Table 1.

5.3 Implications

One of the theoretical implications of this study is that it investigated both the individual and interaction effect of tactile interaction and movement orientation on users’ task performance and subjective experiences. Previous studies comparing the effect of scrolling and swiping used asymmetrical designs, in which they either compared vertical scrolling with horizontal swiping (Kim et al., 2016), or a one-step swipe with a two-step scroll and tap (Choi et al., 2016). This study used a 2 (tactile interaction: scrolling vs. swiping) by 2 (movement orientation: horizontal vs. vertical) cleaner design to tease apart these effects.

Another implication is that the different cognitive load placed by scrolling and swiping mode due to spatial encoding might interact with the type of information. Previous studies that applied cognitive load theory to scrolling and swiping interfaces used textual information (e.g. articles and textual search results) as stimuli and found user performed better using swiping interfaces compared with scrolling interfaces (Kim et al., 2016; Wang & Yueh, 2014; Sanchez & Wiley, 2009; Piolat et al., 1997). This study used pictorial information (e.g. product pictures) and found opposite patterns. The workload and time for judging whether paragraphs of texts or images match a target are different. For pictures, people can scroll and recognize them faster, but for text, people have to finish scrolling and then read through the text to decide whether it's the target. The results of this study suggested that cognitive load might not differ that much between scrolling and swiping interfaces if using pictures as stimuli. Future research aims to investigate different loads placed by different tactile interactions should also take the type of information/stimuli into account.

The current study offers several practical implications for human-computer interaction. Based on the results of this study, I suggested that when designing a mobile search interface for pictures (e.g. shopping interface, visual search result pages), considering using scrolling mode instead of swiping mode if the interface aim to achieve better search performance and the same level of user satisfaction. The reason that I emphasized visual search interface is that previous literature about text search or reading generated opposite results, in which swiping mode generate better task performances than scrolling mode (Kim et al., 2016; Piolat et al., 1997).

In addition, I suggest that web/app designer add some visual cues (e.g. an arrow indicate the direction of scroll) right on the interface. My results showed that some participant would just ignore the instruction, even though I added vivid visual examples and had them stayed on the instruction page for at least ten seconds, they still missed the important information about how to navigate through a list of products on the interface. This is similar to how users skip/ignore tutorials when they first use an app. Designers cannot assume every user would read the tutorial pages carefully and get to know how to use the interface. Adding direct visual cues on top of the interface in addition to tutorial pages would be helpful when introducing new features or helping a new user get familiar with an interface.

REFERENCES

Attfield, S., Kazai, G., Lalmas, M., & Piwowarski, B. (2011, February). Towards a science of user engagement (position paper). In WSDM workshop on user modelling for Web applications (pp. 9-12).

Ayres, P. L. (2001). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology*, 26(2), 227-248.

Baccino, T., & Pynte, J. (1994). Spatial coding and discourse models during text reading. *Language and Cognitive Processes*, 9(2), 143-155.

Baldauf, D., Burgard, E., & Wittmann, M. (2009). Time perception as a workload measure in simulated car driving. *Applied Ergonomics*, 40(5), 929-935.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language,* 68(3), 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Batho, L. P., Martinussen, R., & Wiener, J. (2020). The effects of different types of environmental noise on academic performance and perceived task difficulty in adolescents with ADHD. *Journal of Attention Disorders*, 24(8), 1181-1191.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. Handbook of psychophysiology, 2(142-162).

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hypothesis. Journal of Experimental psychology: Human Learning and memory, 4(6), 656.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica*, 134(3), 330-343.

Borsci, S., Federici, S., Mele, M. L., & Conti, M. (2015, August). Short scales of satisfaction assessment: A proxy to involve disabled users in the usability testing of websites. In International Conference on Human-Computer Interaction (pp. 35-42). Springer, Cham.

Braganza, C., Marriott, K., Moulder, P., Wybrow, M., & Dwyer, T. (2009, April). Scrolling behaviour with single-and multi-column layout. In Proceedings of the 18th international conference on World wide web (pp. 831-840).

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 53-61.

Buetti, S., & Lleras, A. (2012). Perceiving control over aversive and fearful events can alter how we experience those events: an investigation of time perception in spider-fearful individuals. *Frontiers in Psychology*, 3, 337.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology*, 10(2), 151-170.

Chen, S., Epps, J., & Chen, F. (2011, November). A comparison of four methods for cognitive load measurement. In Proceedings of the 23rd Australian Computer-Human Interaction Conference (pp. 76-79).

Chinchanachokchai, S., Duff, B. R., & Sar, S. (2015). The effect of multitasking on time perception, enjoyment, and ad evaluation. *Computers in Human Behavior*, 45, 185-191.

Choi, B. C., Kirshner, S. N., & Wu, Y. (2016, July). Swiping vs. scrolling in mobile shopping applications. In International Conference on HCI in Business, Government, and Organizations (pp. 177-188). Springer, Cham.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load?. *Computers in Human Behavior*, 25(2), 315-324.

Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, 291(5509), 1803-1806.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, 38(2), 105-134.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, 100(1), 223.

Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22(5), 323-327.

Grondin, S., & Plourde, M. (2007). Judging multi-minute intervals retrospectively. *Quarterly Journal of Experimental Psychology*, 60(9), 1303-1312.

Hamilton, P. (1979). Process Entropy and Cognitive Control: Mental Load in Internalised Thought Processes. In Mental Workload (pp. 289-297). Springer, Boston, MA.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in psychology (Vol. 52, pp. 139-183). North-Holland.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need?. *Educational Psychology Review*, 23(1), 1-19.

Kang, H., & Shin, W. (2016). Do smartphone power users protect mobile privacy better than nonpower users? Exploring power usage as a factor in mobile privacy protection and disclosure. *Cyberpsychology, Behavior, and Social Networking*, 19(3), 179-185.

Karray, F., Alemzadeh, M., Abou Saleh, J., & Arab, M. N. (2008). Human-computer interaction: Overview on state of the art.

Khan, A., Sharma, N. K., & Dixit, S. (2006). Effect of cognitive load and paradigm on time perception. *Journal of the Indian Academy of Applied Psychology*, 32(1), 37-42.

Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., & Yoon, H. J. (2016, October). Pagination versus scrolling in mobile web search. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (pp. 751-760).

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46-70.

Lang, A., & Basil, M. D. (1998). Attention, resource allocation, and communication research: What do secondary task reaction times measure, anyway?. Annals of the International Communication Association, 21(1), 443-458.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive sciences*, 9(2), 75-82.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, 19(3), 143-148.

Lewis, J. R., & Sauro, J. (2009, July). The factor structure of the system usability scale. In International conference on human centered design (pp. 94-103). Springer, Berlin, Heidelberg.

Lewis, J. R., Utesch, B. S., & Maher, D. E. (2013, April). UMUX-LITE: when there's no time for the SUS. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2099-2102).

Logan, G. D. (1978). Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology*: General, 107(1), 32.

Lovelace, E. A., & Southall, S. D. (1983). Memory for words in prose and their locations on the page. *Memory & Cognition*, 11(5), 429-434.

Luo, C., Li, X., Liu, Y., Sakai, T., Zhang, F., Zhang, M., & Ma, S. (2017, March). Investigating users' time perception during web search. In Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval (pp. 127-136).

Marathe, S., Sundar, S. S., Nije Bijvank, M., van Vugt, H. C., & Veldhuis, J. (2007). Who are these power users anyway? Building a psychological profile.

Milenkovic, J. (2020, July 20). How Much Time Does the Average Person Spend on Their Phone? Retrieved November 18, 2020, from https://kommandotech.com/statistics/how-much-time-does-the-average-person-spend-on-their-phone/

Montuschi, P., Sanna, A., Lamberti, F., & Paravati, G. (2014). Human-Computer interaction: Present and future trends. *Computing Now*, 7(9).

Moss, A., & Litman, L. (2018). After the bot scare: Understanding what’s been happening with data collection on MTurk and how to stop it. Retrieved February, 4, 2019.

Novak, T. P., Hoffman, D. L., & Yung, Y. F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science*, 19(1), 22-42.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, 33(1), 17-33.

Oh, J., & Sundar, S. S. (2015). How does interactivity persuade? An experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *Journal of Communication*, 65(2), 213-236.

Oh, J., Robinson, H. R., & Lee, J. Y. (2013). Page flipping vs. clicking: The impact of naturally mapped interaction technique on user learning and attitudes. *Computers in Human Behavior*, 29(4), 1334-1341.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429.

Paas, F. G. W. C., Van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79(1 Pt 2), 419-430.

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*(4), 351-371.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63-71.

Penney, T. B., Gibbon, J., & Meck, W. H. (2008). Categorical scaling of duration bisection in pigeons (Columba livia), mice (Mus musculus), and humans (Homo sapiens). *Psychological Science*, 19(11), 1103-1109.

Piolat, A., Roussey, J. Y., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. *International Journal of Human-computer Studies*, 47(4), 565-589.

Ren, H., Kang, H., Ryu, S., & Han, K. (2017). The Effects of Swiping Orientation on Preference and Willingness to Pay: The Interaction Between Touch Interface and Need-For-Touch. 감성과학, 20(4), 65-78.

Sackett, A. M., Meyvis, T., Nelson, L. D., Converse, B. A., & Sackett, A. L. (2010). You're having fun when time flies: The hedonic consequences of subjective time progression. *Psychological Science*, 21(1), 111-117.

Sanchez, C. A., & Wiley, J. (2009). To scroll or not to scroll: Scrolling, working memory capacity, and comprehending complex texts. *Human Factors*, 51(5), 730-738.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4), 469-508.

Sucala, M., Scheckner, B., & David, D. (2011). Psychological time: interval length judgments and subjective passage of time judgments. Current psychology letters. *Behaviour, brain & cognition*, 26(2, 2010).

Sundar, S. S. (2007). Technology and credibility: Cognitive heuristics cued by modality, agency, interactivity and navigability. *Digital media, youth, and credibility. MacArthur Foundation Series on Digital Media and Learning*, 73-100.

Swanston, M. T., Wade, N. J., & Day, R. H. (1987). The representation of uniform motion in vision. Perception, 16(2), 143-159.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.

Sweller, J. (1999). Instructional design. *In Australian educational review*.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, 10(3), 251-296.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, 41(2), 167-174.

Wang, Y. C., & Yueh, H. P. (2014). The Effect of Length in Employment Sites' Web Form Design on User Preferences. *Journal of Library & Information Studies*, 12(1).

Whelan, R. (2008). Effective analysis of reaction time data. The Psychological Record, 58(3), 475-482.

Xie, B., & Salvendy, G. (2000). Prediction of mental workload in single and multiple tasks environments. *International Journal of Cognitive Ergonomics*, 4(3), 213-242.

Xu, Q., & Sundar, S. S. (2014). Lights, camera, music, interaction! Interactive persuasion in e-commerce. *Communication Research*, 41(2), 282-308.

Xu, S., & David, P. (2018). Distortions in time perceptions during task switching. *Computers in Human Behavior*, 80, 362-369.

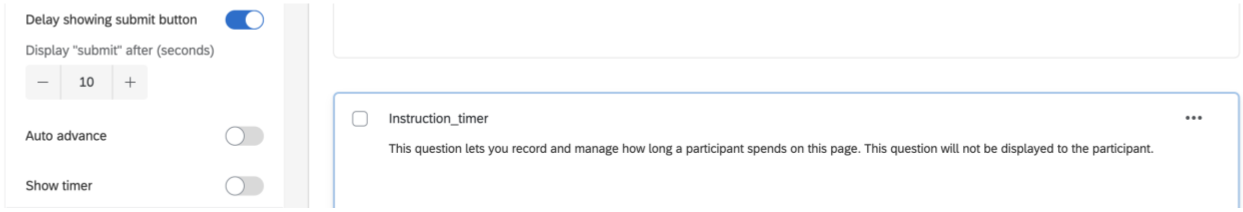
Yao, M. Z., & Ling, R. (2020). "What Is Computer-Mediated Communication?"—An Introduction to the Special Issue. *Journal of Computer-Mediated Communication*, 25(1), 4-8.

APPENDIX

APPENDIX A : INSTRUCTIONS FOR STUDY 1

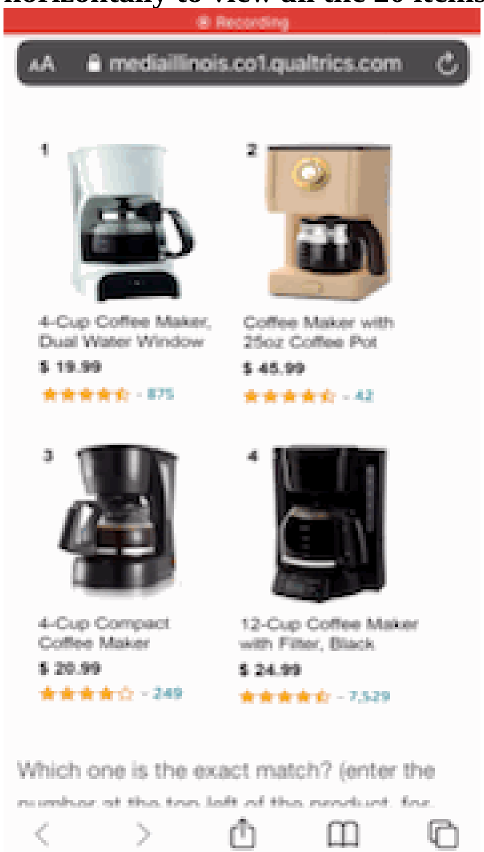
In this study, you will use your mobile phone to complete several search tasks. For each task, **you need to find a specific product in a list of 20 items. You can swipe the list horizontally to view all the 20 items.**

Each question has only one correct answer. Following are two practice trials to get you familiar with the task.



APPENDIX B: INSTRUCTIONS FOR STUDY 2

In this study, you will use your mobile phone to complete several search tasks. For each task, **you need to find a specific product in a list of 20 items. You can swipe the list horizontally to view all the 20 items.**



Each question has only one correct answer. Following are two practice trials to get you familiar with the task.