International Journal of Human-Computer Studies 130 (2019) 130–151

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Post-transitioning user performance on cross-device menu interfaces

Khalid Majrashi

*Department of Information Technology, Institute of Public Administration, 205 Riyadh 11141, Saudi Arabia*

A R T I C L E I N F O

*Keywords:*
Cross-device menus
Redundant menus
Complementary menus
Post-transitioning user performance
Learnability
Inter-usability

A B S T R A C T

Multi-device user-interaction behavior creates a need to design cross-device menus that support users to re-locate menu items efficiently and effectively after the transition from one device to another. We conducted two laboratory experiments to investigate the effects of specific designs of cross-device menus on post-transitioning user performance (selection time and selection error). In Experiment 1, a total of 72 participants moved between 32 desktop and mobile menu interfaces (representing 16 different designs of cross-device menus) to re-locate menu items. We investigated the effect on user performance of the cross-device menu configuration (complementary versus redundant items), menu length (short versus medium), menu item order (consistent versus inconsistent), and menu format (text-only versus text-and-icon items). The results showed significant main effects of configuration, length and item order with large partial eta-squared sizes. Users performed better with redundant menus, short menus, and menus with consistent item order than they did with complementary menus, medium menus, and menus with inconsistent item order. In Experiment 2, 20 participants moved between eight desktop and mobile menu interfaces (representing four different designs of cross-device menus) to re-locate menu items. We focused on the effect on user performance of the menu layout (horizontal-desktop and vertical-mobile versus vertical-desktop and vertical-mobile) at two different lengths (short versus medium). The results showed no significant effect of menu layout, and a significant effect of length with a small partial eta-squared size. Our study provides a foundation to researchers interested in further investigation of cross-device menu designs, and to practitioners in designing cross-device menus that allow more efficient and effective re-locating of menu items when moving between devices.

## 1. Introduction

The use of digital services has been increasingly split over different kinds of interactive devices, including computers, smartphones, tablets, and televisions (Google, 2012; Microsoft, 2013). Users are adopting multiple devices for different types of needs, and the demand for "always-on services" has been growing rapidly, and continues to grow (Forrester Research, 2013; Microsoft, 2013). Cross-platform services (also referred to as "multi-screen services" or "multi-device services") are prevalent among users who own multiple devices (Shin, 2016). A cross-platform service allows access to the digital contents of a specific service across different computing devices. The cross-platform service has been defined as a "set of user interfaces for a single service encompassing two or more computational platforms (hardware and software) for interacting with the service" (Majrashi, 2016). An example of a cross-platform service is the combination of the desktop website and the mobile application for Facebook. Recent research has indicated that almost all users who own multiple devices state that they have been using their devices to interact with cross-platform services to achieve their computing tasks (Majrashi, 2016).

Researchers on human–computer interaction (HCI) have been conducting studies on menu interfaces to understand factors that influence search time and selection accuracy on single menus within individual devices such as laptops, desktops and smartphones (e.g., Byrne et al., 1999; Halverson and Hornof, 2008; Nilsen, 1991; Yang et al., 2017). However, surprisingly, there is a lack of research seeking to investigate factors that affect post-transitioning user performance when switching between multiple menu interfaces across devices. In this study, we use the term "post-transitioning user performance" to refer to user performance (measured as selection time and selection error) after the user has switched to a menu interface on another device to re-locate menu items that had been targeted on a previous menu interface.

The diversity of the characteristics of multi-device systems, and the constraints or requirements of these devices individually can result in having different menu structures across devices. Therefore, we began our research by conducting an informal analysis of the menu interfaces of a set of cross-platform services to identify the different designs of menus used across devices. Across the devices, we identified menu designs with different menu configurations (i.e., different degrees of availability of menu items across devices, referred to as

*E-mail address:* majrashik@ipa.edu.sa.

https://doi.org/10.1016/j.ijhcs.2019.06.001
Received 18 August 2018; Received in revised form 4 June 2019; Accepted 6 June 2019
Available online 06 June 2019
1071-5819/ 

"complementary menus" and "redundant menus"); menu lengths (e.g., short and medium); menu item order (consistent and inconsistent order of items across devices); menu formats (e.g., text-only and text-and-icon); and menu layouts (e.g., horizontal-desktop vertical-mobile and vertical-desktop vertical-mobile). The results of this informal analysis led us to design two experiments to examine the effect of the different elements of menu design (i.e., menu configuration, menu length, menu item order, menu format, and menu layout) on post-transitioning user performance. The first experiment focused on the effects on post-transitioning user performance of menu configuration, menu length, menu item order, and menu format. The second experiment focused on the effects on post-transitioning user performance of menu layout at different menu lengths. Within each experiment, we also investigated the learnability of the different designs of menus, and highlighted the menu designs that had a more efficient learning curve. Our results for the two experiments provide a foundation for future research on cross-device menu interfaces, and a reliable reference for the design of efficient menu interfaces for cross-platform services.

## 2. Background

In this section, we review three research areas related to our study: cross-device interaction, menu interfaces, and learnability.

### 2.1. Cross-device interaction

Cross-device interaction modes have been investigated in different studies. Google (2012) research on multi-screen usage and cross-platform user behavior identifies two main cross-device user-interaction modes: (1) sequential (moving from one device to another at different times); (2) simultaneous (interacting with more than one device at the same time). Google (2012) finds that simultaneous interaction can involve using more than one device at the same time for either a related or an unrelated activity. Similarly, Jokela et al. (2015) examine several cases of multi-device use of the most common devices, including smartphones, computers, tablets, and home media centers, and identify the following main interaction patterns: sequential (changing the device used during the task once or several times in a sequence); resource lending (focusing on a single device but borrowing resources from other devices); related parallel (performing one task using two or more devices); and unrelated parallel (performing several tasks concurrently using two or more devices). To investigate user performance with menu interfaces after transitioning from one device to another, our study relies on the interaction mode in which users change devices sequentially during a task, identified by Google (2012) and Jokela et al. (2015) as a sequential interaction mode.

HCI researchers have also been examining the elements of cross-platform user experience and inter-usability (Denis and Karsenty, 2004; Majrashi et al., 2016a, 2016b, 2017; Wäljas et al., 2010). The key user-experience elements that have been identified include continuity (how seamlessly tasks are migrated across devices); composition (the organization of data and functionality across devices); and inter-device consistency (the consistency of system components across devices). However, no study has investigated the design of cross-device menu interfaces to enhance cross-platform user experience and inter-usability.

The composition of a cross-platform service can be explained in relation to the way devices are organized or how services are delivered across devices (Denis and Karsenty, 2004; Wäljas et al., 2010). Denis and Karsenty (2004) outline three categories of device redundancy with respect to content availability across devices: redundant devices, complementary devices, and exclusive devices. In the category of redundant devices, cross-device user interfaces allow access to the same data and functions. In the category of complementary devices, the cross-device user interfaces share some data and functions, but one (or more) of the devices provides access to data or functions that are inaccessible on the other device(s). In the category of exclusive devices, each cross-device user interface provides access to different data and functions. Denis and Karsenty (2004) find that the category of complementary device is the most common device configuration. Logically, the degree of content availability of a service across devices affects the configuration of menus across devices. For example, menus of services that have a complementary configuration may not have all items presented across the devices because some menu items are associated with content that is not available on a specific device. Given the lack of research investigating user performance on cross-device menu interfaces with different configurations, in this study, we explore the common configurations of menus on cross-platform services, and then examine post-transitioning user performance with these menus, as well as the learnability of the different configurations of menus.

In relation to inter-device consistency, debate continues in the literature about the extent to which inter-device consistency is important for improving the cross-device user experience (Majrashi, 2016). Wäljas et al. (2010) found that consistency may not be a sensitive issue for users, arguing that it may have been overemphasized in many studies on cross-device user experience. Dees (2011) examined the consistency of cross-device user interfaces and states that consistency should not be prioritized over more central inter-usability elements such as task continuity. Pyla et al. (2006) agree with Dees (2011). However, Denis and Karsenty (2004) and Majrashi (2016) suggest that a cross-platform service should maximize consistency to enhance the cross-platform user experience. In addition, Denis and Karsenty (2004) and Majrashi (2016) emphasize the importance of consistency in supporting continuity and seamless user transition between interfaces across devices. In our study, consistency is adopted as an independent variable to investigate whether it plays a role in enhancing post-transitioning user performance on cross-device menu interfaces and the learnability of menus, and in turn supports task continuity and seamless transitioning between devices as ultimate inter-usability goals.

Majrashi (2016) studied user behavior when interacting with services across different devices and identifies several cross-platform user-interaction behavioral patterns. These patterns include visual memory, where users sometimes re-find objects on another device by remembering their appearance (e.g., shapes); habituation, where some of the users' frequent physical actions become reflexive when using the interface on another device; and spatial memory, where users sometimes re-find objects on another device by remembering their location on the interface (e.g., top, bottom). However, no research has specifically examined whether such behaviors influence post-transitioning user performance on cross-device menu interfaces. Thus, we examine and discuss variables relevant to these behaviors, including the use of icons in the menus and the position of items within menus across different devices.

### 2.2. Menu interfaces

Many studies have been conducted to investigate operating performance with menu interfaces. For example, Callahan et al. (1988) examined user performance with linear and pie menus on personal computers and found that pie menus outperform linear menus in relation to efficiency and lowering error rates. Yang et al. (2017) conducted a study on user performance with half-pie and linear menus on smartphones, specifically examining users thumb sizes. A key finding was that the performance of users with shorter thumbs was better than that of users with longer thumbs. Kacmar and Carey (1991) conducted an experiment to measure user performance with menu items constructed of text, icons, and text-and-icons in a personal-computer-based environment, and found that menus constructed of text-and-icons had a lower rate of incorrect selections by users.

Research has also identified different factors affecting user performance when operating menus. For example, efficiency of users is better with menus that have fewer items (Byrne et al., 1999; Halverson and

Hornof, 2008; Nilsen, 1991); selection time is affected by organization of items on the menu (Card, 1982; Halverson and Hornof, 2008; McDonald et al., 1983); search time can be shorter for menu targets that are closer to the top of the menu (Byrne et al., 1999; Cockburn et al., 2007); and second and last items can also have shorter selection times (Bailly et al., 2014; Byrne, 2001; Byrne et al., 1999). While all previous studies have focused on measuring user performance with single menus within the environment of a single device, our study examines user performance when switching between multiple menu interfaces across devices.

### 2.3. Learnability

In 1885, the psychologist Ebbinghaus described the term "learning curve" as referring to task efficiency being improved when the same task is repeated many times (Green, 1999). In HCI and usability, learnability is a common attribute of high quality design. Michelsen et al. (1980) state that a system that has a high level of learnability "should be easy to learn by the class of users for whom it is intended." Nielsen (1994) states that the level of learnability refers to how easy it is for first-time users to accomplish tasks on the system. Learnability can be measured through comparing performance on an initial trial with performance after repeated trials (Sauro, 2013).

The learnability of interface menus has been investigated in many studies (e.g., Cockburn et al., 2007; Kaptelinin, 1993; Somberg, 1987; Yang et al., 2017). The factors of interface menus that affect learnability have also been discussed in the literature. For example, Somberg (1987) explains that users performing repeated tasks can remember the location of items on the menu interface and will look at them immediately if no changes have been made to the menu layout.

In the context of inter-usability research, Majrashi (2016) defines learnability as the "extent to which a cross-platform service (a set of user interfaces) can be learned efficiently," and how easy it is for users to re-accomplish tasks performed on a specific interface after transitioning to another user interface without new learning requirements. One aim of our study is to measure the level of learnability of cross-device menu interfaces by comparing menu operating performance of initial trials with menu operating performance of repeated trials, and analyze the different elements of menu design that affect learnability.

## 3. Experiment 1

This section presents the research methodology, providing a description of the experimental menu-interface design, the participants, and experimental design, as well as of the apparatus, measures and stimulus, and the procedures for Experiment 1. This section also presents the experimental results.

### 3.1. Method

#### 3.1.1. Analysis of common cross-platform-service menu interfaces

The combination of a smartphone and laptop or personal computer is one of the most common combinations of devices employed by many users every day to perform related or unrelated cross-platform computing activities (Google, 2012). Our study employed these combinations of devices to examine cross-platform services and analyze service menus, and then conduct our experiments.

To design navigation menus for our experiments that would be representative of common menu configurations seen in real-world cross-platform services, we conducted an informal analysis of the main navigation menus for 55 popular services. Each service was a combination of a desktop website and mobile application or site. Such analyses are commonly applied in research to understand the characteristics of modern application menus before conducting experiments (see Bailly et al., 2008). Our informal analysis focused on the first-level menu items of browse-mostly services. The services were from different domains, including email, online shopping, travel, social networking, financial management, and video watching, which are popular with users who regularly perform cross-platform activities (Google, 2012). We selected the services from a list of top websites published by Alexa (www.alexa.com) in each domain. After selecting a specific service, statistics of the service's mobile application (e.g., download statistics) were checked to ensure its popularity. Forty-six of the 55 services apply a global-navigation pattern for desktop interfaces and use side-navigation-drawer menus for mobile interfaces. Our analysis was based on these 46 services, with the intention of analyzing and studying the services that apply other types of navigation patterns in future research. Of those 46 services, 26 have a complementary menu configuration and 20 have a redundant menu configuration. It is important to note here that we use the terms "complementary" and "redundant" to describe the configuration of menu items rather than to describe the configuration of the entire content of all user interfaces as do Denis and Karsenty (2004). Thus, in our study, a complementary menu refers to menus on which some items are available across devices but that have one or more item not accessible on the other device, and a redundant menu refers to all menus having the same items.

For each type of menu configuration, our informal analysis identified three categories of menu length across devices: short menus (5–10 items); medium menus (11–15 items); long menus (>15 items). Short and medium menus were the most common categories in the analyzed services, leading us to focus on these lengths of menu in our experiments (see Tables 1 and 2). We designed two complementary menus (short and medium) and two redundant menus (short and medium). For the complementary menus, we configured the length of the desktop and mobile menus based on the average length of the analyzed menus associated with each device (Table 1, column 4). The complementary menus were designed to have a set of shared items between the desktop and mobile menus. The number of shared items was determined based on the average number of shared items identified per menu-length category (Table 1, column 5).

For the redundant menus, we configured the menus depending on the average length of the menus (seven items for the short menus and 12 items for the medium menus) (Table 2, column 3). While a redundant menu refers to cross-device menus that have the same items, the length of a specific redundant menu across desktop and mobile devices is always the same.

**Table 1**
Results of informal analysis of complementary menus in popular cross-platform services.

| Category of menu | Percent of services | Device(s) | Menu length (average number of items) | Shared items (average number of items) |
|---|---|---|---|---|
| Short (5–10) | 77 | Desktop | 8 | 5 |
| | | Mobile | 7 | |
| Medium (11–15) | 15 | Desktop | 12 | 10 |
| | | Mobile | 13 | |
| Long (>15) | 8 | Desktop | 19 | 13 |
| | | Mobile | 28 | |

Percentages rounded to the nearest whole number.

**Table 2**
Results of informal analysis of redundant menus in popular cross-platform services across desktop and mobile devices.

| Category of menu | Percent of services | Menu length (average number of items) |
|---|---|---|
| Short (5–10) | 75 | 7 |
| Medium (11–15) | 20 | 12 |
| Long (>15) | 5 | 17 |

Percentages rounded to the nearest whole number.

In our informal analysis, we also examined the consistency of item order across devices. We found that 74% of redundant menus had a consistent item order and that only 8% of complementary menus had a consistent item order starting from the top of the menus. This led us to adopt consistency of item order as an experimental condition and perform our experiments with menu items organized consistently and inconsistently, with the aim of comparing user performance between consistent and inconsistent menus. Our analysis also found that 73% of corresponding shared items in complementary menus and 95% of corresponding shared items in redundant menus used consistent labels. This led us to design menus with consistent labels, with the intention of investigating inconsistent labeling of cross-device menus in future research. In relation to using text in combination with icons in menus, we found that approximately 20% of analyzed menus used text-and-icon (with an icon beside the text for the item) across the devices. We adopted these results and used menu format (i.e., text-only versus text-and-icon) as another experimental condition. Based on our analysis of the design of cross-platform service menu interfaces, we summarize the independent variables with their levels for Experiment 1 in Table 3.

#### 3.1.2. Menu design

We designed Arabic-menu interfaces because our study was conducted in Saudi Arabia. We used 45 unique Arabic single-word items across all menu configurations. The set of words in each menu was selected from real-world applications for each specific domain associated with the menu. The short-complementary menu (SC) contained words from a furniture shopping service and the medium-complementary menu (MC) contained words from a newspaper service. The short-redundant menu (SR) contained words from a clothing shopping service and the medium-redundant menu (MR) contained words from an online sports-news service. Each menu had words with no more than three letters difference in length to reduce any possible saliency effects (Bailly et al., 2014). We did not use the same words across the different menus to reduce any potential of confusing participants given that they were required to interact with the different menus.

We used the Arabic word for "Home" in the first position of each menu to ensure consistency with most of the analyzed menus and to simulate real-world menus. If we did not use the word "Home" in the first position, it is likely that users (based on their previous experiences and expectation) would ignore the item placed in the first position because it is normally reserved for "Home".

For the desktop interface, the horizontal separation between items was 50 pixels. The menu items were 80 pixels from the top of the screen, and the first item was 200 pixels from the right of the screen. The desktop menu items were presented in a "Simplified Arabic" 14-point font (~4.91 mm). The selection of this combination of font type and size was based on previous studies indicating that this combination eases the digital and online reading process for Arabic users (Alsumait et al., 2009; Ramadan, 2011).

For the mobile interface, we designed the side navigation drawer following design recommendations by Google (2018). However, we modified the direction of elements to suit Arabic as a language that reads from right to left. The text of the menu items was aligned right, and the right horizontal margin was 16 pixels from the edge of the navigation drawer. The first item was 40 pixels from the top of the screen. The vertical spacing between menu items was 42 pixels. The physical target size for menu items was at least 9 mm in height and width to ensure that they would be easy for users to tap, thus preventing unintentional selection of menu items. We used the same font type used for the desktop menu but increased the font size to 15 points (~5.27 mm) to improve the readability of the Arabic font on the small screen (Google, 2018). We used a horizontal line (separator) between the items on the mobile menus, which is consistent with the design of approximately 60% of the mobile menus we analyzed.

The mobile menu was designed to be visible when loading the mobile interface, meaning that users did not need to click on the menu icon to view the menu. All items in each mobile menu were displayed in a single view at the same time so that users did not need to scroll down to find more items. These adjustments were crucial for reducing any effects resulting from the difficulty in finding a hidden navigation menu (Kara and Raluca, 2016) and from the difficulty in scrolling (Chae and Kim, 2004), and therefore for the accuracy of our measurements.

For the menus with text-and-icon format, we placed the icons before the text. We considered the timing for movement from the initial position (the text link invoking the menus) to the target item once the target is located, following Fitts' law (MacKenzie, 1992). Thus, we ensured the distance from the initial position to the menu items in the text-only menus was equal to the distance from the initial position to the icons in the text-and-icon menus. The icons had equal width and height (24 × 24 pixels) in the mobile and desktop interface. The icons were adopted from real-world applications and online icon libraries. Five participants were asked to match icons and texts to ensure that each icon could be interpreted by users as representing only one menu item, and to assess the relatedness between the icons and the text for the menu items. In this test, the icons and texts were presented in two facing columns and shuffled randomly, and respondents were instructed to skip any icon that had no meaningful association with any text. Participants rated the relatedness between the icons and texts using a seven-point Likert-type scale (1 = very unrelated, 7 = very related). The results showed that all participants produced the same associations as our intended associations, and each icon and text pair had an average relatedness equal to or above 6.4. Thus, we reduced the likelihood of a confounding effect resulting from misinterpretation of icon meanings and unrelatedness between the icons and labels.

For the mobile interface, the icons were aligned right, and the right horizontal margin was 16 pixels from the edge of the navigation

**Table 3**
Independent variables and their levels in Experiment 1.

| Independent variable | Levels | Description |
|---|---|---|
| Menu configuration | Complementary | Desktop and mobile menus have a set of shared items with exclusive items per menu |
| | Redundant | Desktop and mobile menus have the same items |
| Menu length | Short | Desktop and mobile menus have a short length |
| | Medium | Desktop and mobile menus have a medium length |
| Menu item order | Consistent | Desktop and mobile menus have a consistent order of items |
| | Inconsistent | Desktop and mobile menus have an inconsistent order of items |
| Menu format | Text-only | Desktop and mobile menus use text-only for items |
| | Text-and-icon | Desktop and mobile menus use text-and-icon for items |

**Table 4**
Design of Experiment 1.

| | | | | Group | Menu configuration<br>Complementary<br>Menu length<br>Short | <br><br><br>Medium | <br>Redundant<br><br>Short | <br><br><br>Medium |
|---|---|---|---|---|---|---|---|---|
| Menu format | Text-only | Menu item order | Inconsistent | A | SC | MC | SR | MR |
| | | | Consistent | B | SC | MC | SR | MR |
| | Text-and-icon | | Inconsistent | C | SC | MC | SR | MR |
| | | | Consistent | D | SC | MC | SR | MR |

SC = short-complementary, MC = medium-complementary, SR = short-redundant, MR = medium-redundant.

drawer. The space between the icons and the text of the menu item was 56 pixels.

In total, we had 32 desktop and mobile menu interfaces, representing 16 different designs of cross-device menu interfaces (see Table 4). illustrates text-only and text-and-icon consistent short-redundant menus and medium-redundant menus as examples of the menu interfaces in Experiment 1.

### 3.1.3. Participants and experimental design

Seventy-two graduate students (36 females and 36 males), aged 25 to 29 years with more than two years of internet and cross-device experience participated in the study voluntarily. All participants had a bachelor's degree as their highest education level. Participants did not receive any form of compensation for taking part in this experiment. All were native Arabic speakers and had normal vision. Participants were divided randomly into four main groups (A, B, C, and D) with 18 students in each group. Each group operated eight desktop and mobile menus, representing four different cross-device menu configurations and lengths. We used a mixed design of 2 (menu configuration: complementary and redundant) × 2 (menu length: short and medium) × 2 (menu format: text-only and text-and-icon) and × 2 (menu item order: consistent and inconsistent) (see Table 4). The variables of menu format and menu item order were manipulated between subjects, while the variables of menu configuration and menu length were manipulated within subjects. The dependent variables were post-transitioning selection time and selection error. Post-transitioning selection time refers to the time spent by a user on a task when they had switched to another device to re-locate a menu target. Post-transitioning selection error refers to the error(s) made by a user when performing a task when they had switched to another device to re-locate a menu target.

For the between-subjects design, the controlled variables were the devices' specification and menu items. The menu item order for the variable of menu format (text-only versus text-and-icon) was controlled. We also maintained target positions across the experimental conditions (consistent versus inconsistent item order) to eliminate target-position effects. That is, users in each experimental condition searched for items located in pre-specified positions. For example, if users searched items b, c, d, and e located in positions 2, 3, 4, and 5 in the first condition (consistent item order), in the second condition (inconsistent item order), the participants searched the same items (b, c, d, and e) by arranging them in the same positions (2, 3, 4, and 5) but in a different order across devices. The variable of device specification was controlled for the within-subjects design.

The tasks for the different menus were generated randomly before the experiment began. The same tasks were then adopted by the system for the different groups of participants across the different conditions. Thus, the tasks were controlled for the between-subjects design. The item order for the redundant menus was fully consistent across the devices in the consistent condition. The item order of the shared items in the complementary menus was also fully consistent across the devices, starting from the top of the menus, followed by the non-shared items at the bottom of the menus. There was a 100% level of inconsistency of item order in the inconsistent condition for all menus, meaning that the corresponding target items were always in different positions across the devices. We did this to ensure that our measurements were based on the selection of items placed in inconsistent positions. In the complementary menus, participants always searched items from the shared menu items across devices because our primary aim in this study is to measure user performance when re-locating items after transitioning from one device to another.

When assigning participants to the four groups, we made our best attempt to counterbalance confounding factors such as age and gender. We randomly assigned nine females and nine males to each group, although we found that gender did not play a significant role in the study results. As reported by Dillon and Watson (1996), Egan (1988), and Lazar et al. (2010), individual differences are expected to be small and have a low effect on results in research similar to the present study because the type of task (target selection) is simple and involves limited cognitive processes.

Each participant operated the menus in four sessions, one session for each of the combinations of menu configuration and length: short-complementary, medium-complementary, short-redundant, and medium-redundant. The order of the sessions was counterbalanced within each group of participants. Each participant conducted 38 cross-device search trials, divided into four blocks (the number of tasks in each block depended on the menu configuration and length). In each trial, each participant attempted two tasks (one with each device). To investigate the learning curve of menus, participants operated each menu three times. In summary, we had 72 participants × 38 trials × 2 tasks per trial × 3 operations of each menu = 16,416 selections. Half of these selections (8208) were associated with re-location of menu items after switching from one device to another. Our analysis of post-transitioning user performance was based on the selections associated with re-location of menu items. The entire experiment required approximately 50 min to complete for each user.

### 3.1.4. Apparatus, measures and stimulus

Our experiment used MacBook Pro (13 inch) and Samsung Galaxy S9 smartphone. The specifications of MacBook Pro are 2.3 GHz dual-core Intel Core i5; dimensions 30.41 cm × 21.24 cm; resolution 2560 × 1600 at 227 pixels per inch; macOS High Sierra; and 8 GB RAM. The smartphone specifications are dimensions 147.7 × 68.7 × 8.5 mm (5.81 × 2.70 × 0.33 inches); weight 163 g (5.75 oz); size 5.8 inches, 84.8 $cm^2$ (~83.6% screen-to-body ratio); resolution 1440 × 2960 pixels; 18.5:9 ratio (~570 ppi density); Android 8.0 (Oreo) OS; Octa-core 4 × 2.7 GHz CPU; and 4 GB RAM.

We used two measures: selection time and selection error. Selection time was recorded from the moment the menu was displayed until the participant selected the correct target. The time was recorded automatically by the prototype systems. Selection error was recorded by the prototype systems when an incorrect selection of the targeted item occurred. The systems also recorded the number of incorrect selections per task for each participant. Our study focused on the post-transitioning selection time and post-transitioning selection error.

As in many related studies (e.g., Bailly et al., 2008, 2014), we used the explicit target name of the item as the stimulus for all menus in our


**Fig. 1.** Illustration of text-only and text-and-icon consistent short-redundant menus and consistent medium-redundant menus in English for desktop and mobile interfaces.

experiments. This was to avoid any possible confounding effects that could result from using other types of stimuli such as an icon, synonym of target label, or definition of target. The stimulus (the exact label for the target item) was presented in a text window at top of the menu. For the stimulus (target name of the item), we used the font type and size used for menu items in each menu interface, "Simplified Arabic" 14-point font for stimulus in the desktop interfaces and "Simplified Arabic" 15-point font for stimulus in the mobile interfaces. This was to eliminate any possible saliency-affecting factors that may result from using inconsistent font type and size for stimuli and menu items.

*3.1.5. Procedure*

At the beginning of the experiments, participants were asked to answer a set of questions about their background information and internet and device experience. We then informed participants that they would be required to search menus to locate target items. The participants were told that an explicit target name for the item that needed to be located would be displayed on the screen at the beginning of each

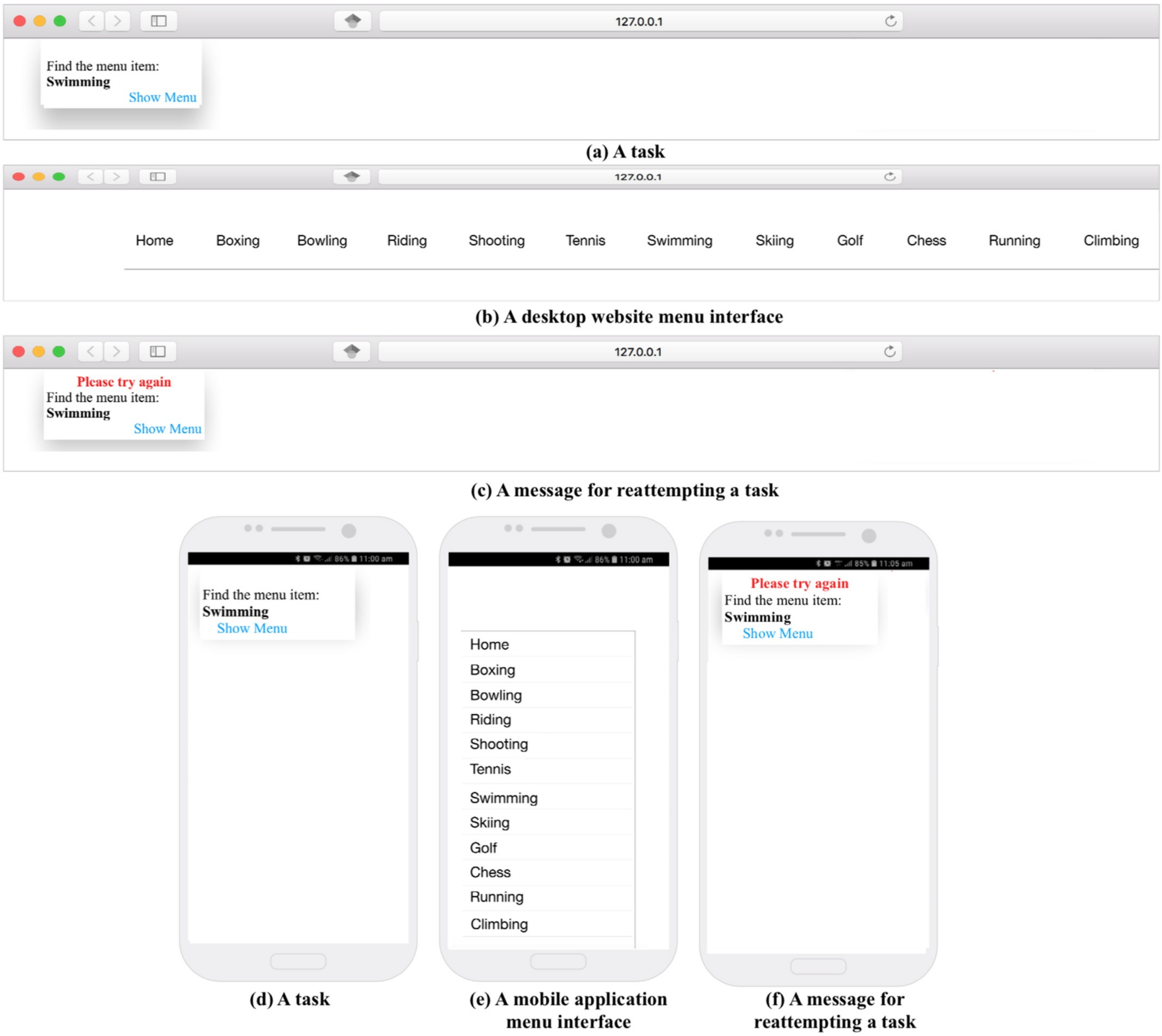



**Fig. 2.** Illustration of experimental desktop (a, b, and c) and mobile (d, e and f) interfaces (in English).

search trial. The participants were told that the goal was to investigate how quickly they could re-locate a given target after switching from one device to another. Each trial involved searching a targeted item across two devices. Participants were informed that there was only a single target item for each task in each search trial.

At the beginning of the experiments, each participant conducted three practice search trials using desktop and mobile menus, which were different from the menus designed for the experiments. These search trials provided the participants with some time to relax and become comfortable with the devices and the experimental environment. The participants were not told that the first three search trials were not scored. Under each condition, the three practice search trials were conducted on menus similar to those used by the groups under the condition to which they would be assigned to avoid the potential learning effect within a specific group and not within another group. For example, for participants who had been assigned to the text-only consistent menus, we used trial menus designed with a text-only format and a consistent item order, while for participants who had been assigned to the text-and-icon consistent menus, we used trial menus with a text-and-icon format and a consistent item order. The experiment began once the participant felt comfortable in the experimental environment.

The cross-device trials were conducted with the devices in quick succession (i.e., the users switched devices within a short period), which is a common form of cross-device interaction (Majrashi, 2016). The cross-device trial began on the first device by showing the participant a text window containing the stimulus at top of the screen and a text link labeled as “Show Menu” (e.g., Fig. 2a). This represents the first task in the trial. After reading the stimulus, the participant was required to press the text link that was located at the top of the first item of the menu; pressing this text link made the description of the target disappear and the menu appear (Fig. 2b). In similar experiments, it is a common approach to have the starting location at the top of the screen (see Bailly et al., 2014; Byrne, 2001). The first task began as soon as the participant pressed this text link. The first task ended when the participant had selected the correct target item and the menu disappeared. In the case of an incorrect selection, the menus disappeared, and the participant was requested to make another selection by showing the stimuli again and a message asking the participant to reattempt the task (Fig. 2c). Again, the participant was required to press the text link to display the menu. If the participant succeeded in locating the item on the first device, a text window appeared to ask the participant to switch to the second device. On the second device, the participant followed the same procedure to complete the second task in the trial (Fig. 2d, 2e, and 2f). The completion of one trial across devices immediately began the next trial, but with a reverse device order. That is, after a participant had made the correct selection of the item in the second task in the trial, they were shown a text window asking them to wait for another task to be performed on the device they were currently using, and upon completion of that task, another text window appeared to instruct them to switch to the other device. During the experiment, participants were not given feedback on their performance.

NASA Task Load Index (NASA-TLX) has been successfully used as a self-report measure of cognitive workload in usability studies (Als et al., 2005; Hornbæk, 2006; Pasqualotto et al., 2011; Tracy and Albers, 2006). The NASA-TLX is a subjective-measurement technique that rates perceived workload to assess a task, system, or other aspects of performance (Hart and Staveland, 1988). It has six subscales: mental demand (i.e., level of required mental and perceptual activity); physical demand (i.e., level of required physical activity); temporal demand (i.e., time pressure in relation to the task); performance (i.e., level of success in performing the task or level of satisfaction about the performance); effort (i.e., degree of effort to work mentally and physically to accomplish a level of performance); and frustration (i.e., level of frustration during the task). After completing the trials with each tested menu, we asked participants to complete the questions on the NASA-TLX. The questions of the NASA-TLX were designed to focus on measuring the cognitive workload of re-locating menu items. Our study used seven levels (from 1 to 7) of workload for the six subscales of NASA-TLX.

Ethics procedures were followed in compliance with the ethical standards in Saudi Arabia.

### 3.2. Analysis and results

In this section, we analyze our experimental results of the effects on post-transitioning selection time and selection error of menu items of menu configuration (redundant and complementary), menu length (short and medium), menu item order (consistent and inconsistent), and menu format (text-only and text-and-icon). We also analyze the results of the NASA-TLX, and the learning curve of the menus. We used IBM's Statistical Package for the Social Sciences (SPSS) statistical software for analyzing the results.

#### 3.2.1. Post-transitioning selection time

Mixed analysis of variance (ANOVA) with a significance level of 0.05 was used to analyze average post-transitioning selection time and judge the significance of effects. Table 5 presents the average selection times and standard deviation for all menus tested in the experiment. Fig. 3 shows the boxplot of selection time for all tested menus. The overall results showed that users recorded the fastest selection time when operating the text-only consistent short-redundant menu ($M$ = 1.47 s, $SD$ = 0.488), compared with all other tested menus. Users had the slowest selection time with the text-only inconsistent medium-complementary menu ($M$ = 4.49 s, $SD$ = 0.593). A general observation seen in Table 5 is that users had lower selection times when operating short menus than they did when operating medium menus. The differences in selection times between short and medium menus within each test condition are clear from the boxplots in Fig. 3. In all cases, the median of the selection times for the medium menus are larger than the median of the selection times for the short menus. Similar results are observed when comparing the consistent menus and inconsistent menus, in that users generally had faster selection times for the consistent menus.

Another interesting observation is that users had longer selection times for the inconsistent short menus than they did for the consistent medium menus. For example, users took an average of 2.81 s to select a menu target with the text-only inconsistent short-complementary menu compared with 2.04 s with the text-only consistent medium-complementary menu. Another important result is that users recorded faster selection times for the redundant menus than they did for the complementary menus when comparing menus of the same length (short-complementary versus short-redundant and medium-complementary versus medium-redundant), with a single exception, in which the users took 2.38 s to locate the menu target in the text-and-icon inconsistent short-redundant menu and 2.16 s to locate the menu target in the text-and-icon inconsistent short-complementary menu.

Table 6 shows the main and interaction effects of the independent variables on post-transitioning selection time. Effect size was calculated using partial eta-squared ($\eta_p^2$) and interpreted as small ($\geq$0.01), moderate ($\geq$0.06), or large ($\geq$0.14) (Cohen, 1988). The ANOVA test results analysis found a significant effect of configuration, and the effect size was large ($\eta_p^2$ = 0.632). The results of menu configuration (complementary versus redundant) demonstrate that regardless of menu length, users located the targets faster by 0.28 s when operating redundant menus ($M$ = 2.37 s) compared with complementary menus ($M$ = 2.65 s). Fig. 4a illustrates that the median of the selection time of the complementary menus is larger than the median of the selection times of the redundant menus.

**Table 5**
Average post-transitioning selection time and SD.

| Menu format | Menu item order | Menu configuration | Menu length | Average selection time (time taken in seconds per an item selection) | SD |
|---|---|---|---|---|---|
| Text-only | Inconsistent | Complementary | Short | 2.81 | 0.588 |
| | | | Medium | 4.49 | 0.593 |
| | | Redundant | Short | 2.51 | 0.340 |
| | | | Medium | 3.67 | 0.426 |
| | Consistent | Complementary | Short | 1.51 | 0.627 |
| | | | Medium | 2.04 | 0.564 |
| | | Redundant | Short | 1.47 | 0.488 |
| | | | Medium | 1.86 | 0.374 |
| Text-and-icon | Inconsistent | Complementary | Short | 2.16 | 0.547 |
| | | | Medium | 3.98 | 0.731 |
| | | Redundant | Short | 2.38 | 0.331 |
| | | | Medium | 3.66 | 0.619 |
| | Consistent | Complementary | Short | 1.79 | 0.555 |
| | | | Medium | 2.46 | 0.646 |
| | | Redundant | Short | 1.54 | 0.436 |
| | | | Medium | 1.88 | 0.464 |

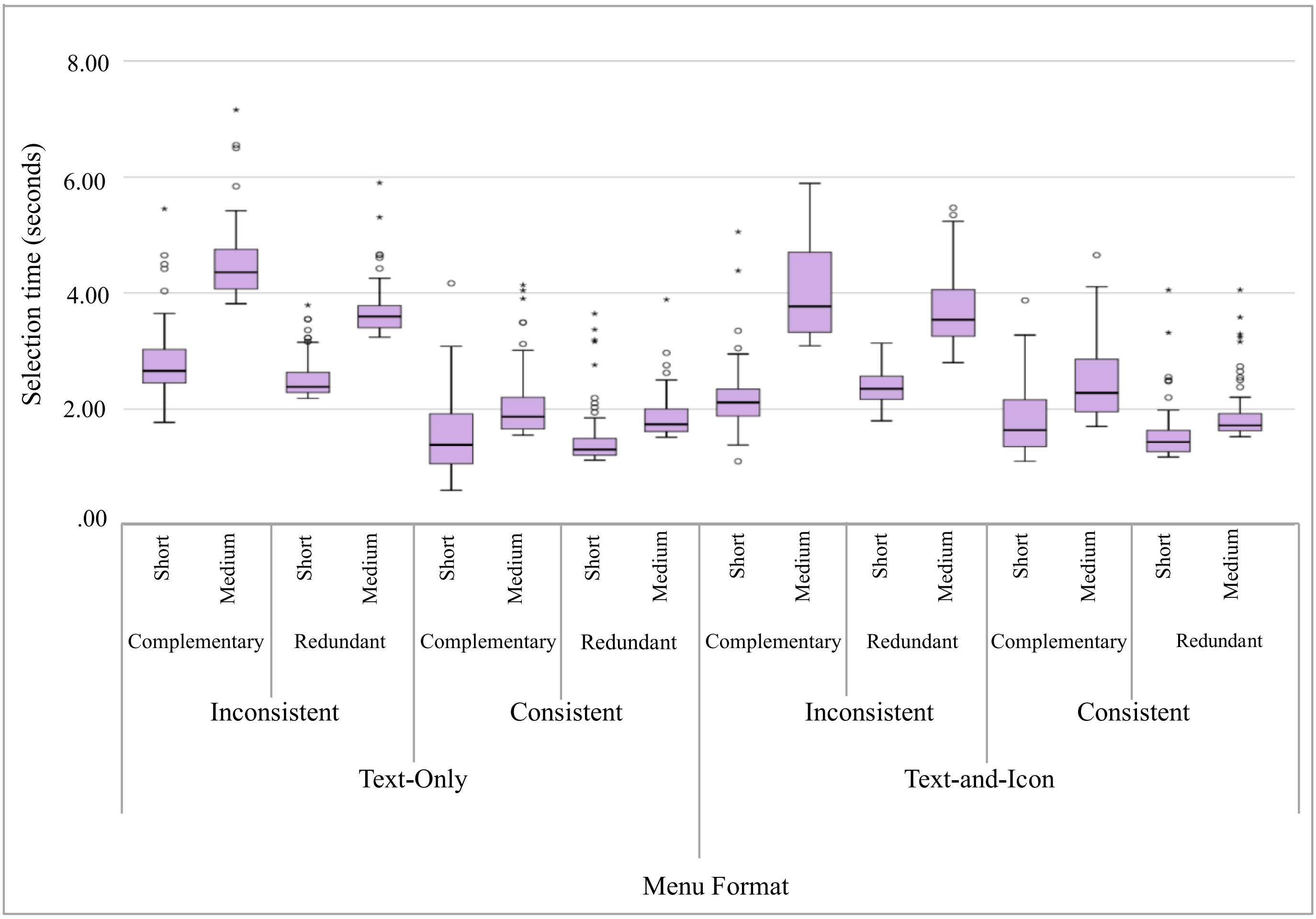



**Fig. 3.** Boxplot of post-transitioning selection time in seconds for all menus in Experiment 1.

**Table 6**
Main and interaction effects of menu configuration, menu length, menu item order, and menu format variables on post-transitioning selection time.

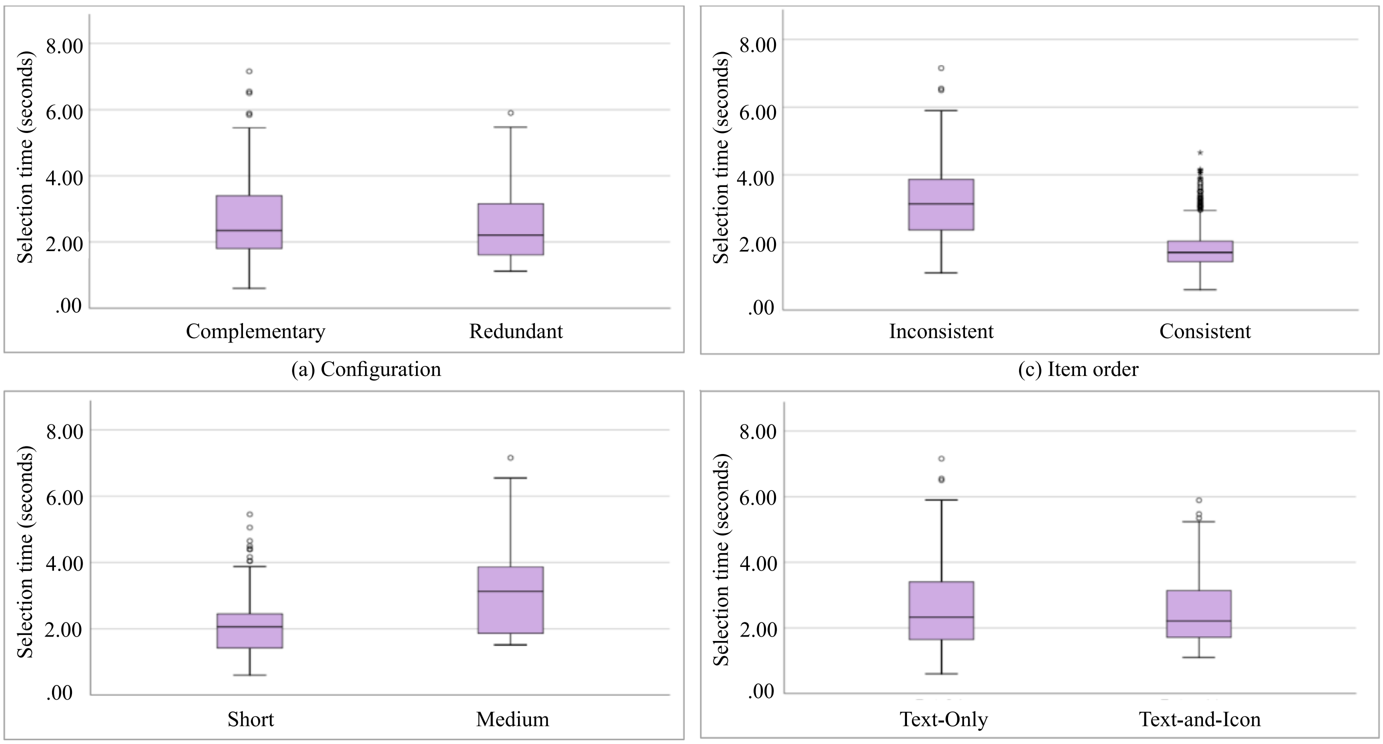
| Variable(s) | $F$ | $P$ | $\eta_p^2$ |
|---|---|---|---|
| Configuration | 610.65 | $<0.001$ | 0.632 |
| Length | 12,333.68 | $<0.001$ | 0.972 |
| Item order | 669.66 | $<0.001$ | 0.653 |
| Menu format | 1.38 | $=0.240$ | 0.004 |
| Configuration × length | 589.53 | $<0.001$ | 0.623 |
| Configuration × item order | 3.19 | $=0.075$ | 0.009 |
| Configuration × menu format | 18.72 | $<0.001$ | 0.050 |
| Length × item order | 3186.88 | $<0.001$ | 0.900 |
| Length × menu format | 25.24 | $<0.001$ | 0.066 |
| Item order × menu format | 23.58 | $<0.001$ | 0.062 |
| Configuration × length × item order | 90.96 | $<0.001$ | 0.204 |
| Configuration × length × menu format | 12.56 | $<0.001$ | 0.034 |
| Configuration × item order × menu format | 322.63 | $<0.001$ | 0.475 |
| Length × item order × menu format | 7.58 | $<0.01$ | 0.021 |
| Configuration × length × item order × menu format | 7.48 | $<0.01$ | 0.021 |

Hypothesis df = 1; error df = 356.

Configuration did not interact significantly with menu item order, suggesting that the effect of configuration does not depend on the consistency of item order across devices. That is, the variance in selection times when comparing inconsistent redundant and inconsistent complementary menus, and consistent redundant and consistent complementary menus are similar. Users recorded 0.30 s faster selection time with inconsistent redundant menus compared with inconsistent complementary menus and 0.26 s faster selection time with consistent redundant menus compared with consistent complementary menus.

There was a significant interaction effect of configuration × menu format. This effect means that there was a difference in the level of the configuration effect on selection time when operating the different menu formats regardless of menu length and menu item order, demonstrating a larger effect of configuration when operating text-and-icon menus. That is, the difference in selection time between complementary and redundant menus is greater in the text-and-icon format than in the text-only format. This was seen in the fact that users took 0.70 s longer to operate complementary menus than redundant menus in the text-and-icon menu format, and 0.33 s longer in the text-only menu format. However, the effect size of configuration × menu format was small ($\eta_p^2 = 0.050$).

A three-way significant interaction effect of configuration × item order × menu format with a large effect size ($\eta_p^2 = 0.475$) was identified. This effect was interpreted as meaning that the complementary menus had a longer selection time than the redundant menus, and the level of difference in selection time depended on menu item order and menu format. Our in-depth investigation of this effect showed that the there is a larger effect of configuration when users searched text-only inconsistent menus, followed by text-and-icon consistent menus, text-only consistent menus, and text-and-icon inconsistent menus. The variance between selection times of complementary menus and redundant menus was 0.56 s for text-only inconsistent menus, 0.41 s for text-and-icon consistent menus, 0.11 s for text-only consistent menus, and 0.05 s for text-and-icon inconsistent menus. The three-way effect of configuration × item order × menu format provided more precise and




(a) Configuration (c) Item order

(b) Length (d) Menu format

**Fig. 4.** Boxplot of post-transitioning selection time in seconds for the levels in each independent variable in Experiment 1.

interpretable results than the two-way effect of configuration × menu format. For example, one interpretation of the three-way effect result is that using icons can contribute to reducing the effect of a complementary menu configuration when operating menus that have an inconsistent item order.

The ANOVA test results showed that selection time is significantly affected by menu length (see Table 6). For the complementary and redundant menu configurations, participants recorded faster selection times with the short menus (short-complementary and short-redundant) than they did with the medium menus (medium-complementary and medium-redundant) (see Table 5). Fig. 4b illustrates that the median of the selection times of medium menus is a great deal larger than the median of the selection times of short menus. The users took approximately 1.00 s longer with medium menus ($M$ = 3.00 s) than with short menus ($M$ = 2.02 s), averaged over all the independent variables. There was also a significant two-way interaction effect of menu length and menu item order, demonstrating that the effect of length on selection time is different when comparing consistent menus with inconsistent menus. As seen in Fig. 3, length had a smaller effect when users were operating consistent menus. Averaged over the variables of menu format, users took 1.48 s longer with the medium menus compared with the short menus in the inconsistent condition. However, users took only 0.48 s longer with medium menus compared with the short menus in the consistent condition. An important observation from Table 6 is that the size of the main effect of length ($\eta_p^2$ = 0.972) and the interaction effect of length × item order ($\eta_p^2$ = 0.900) is very large, representing the largest effect on selection time among all investigated variables.

The interaction between menu length and menu format was also significant. This result demonstrates that the effect of length on selection time is smaller for text-only menus than it is for text-and-icon menus. However, the size of this effect was moderate ($\eta_p^2$ = 0.066). That is, the differences in selection times between short and medium menus in the two conditions are very similar. The users took 0.94 s longer with text-only medium menus compared with text-only short menus and 1.02 s longer with text-and-icon medium menus compared with text-and-icon short menus.

The interaction of configuration × length was also significant, demonstrating that the level of the effect of length can also depend on the type of menu configuration (complementary versus redundant). The size of this effect is large ($\eta_p^2$ = 0.623). This result demonstrates that length has a larger effect on selection time when users operate complementary menus. Averaged over all conditions, users took 1.17 s longer with medium-complementary menus than they did with short-complementary menus. However, they took 0.79 s longer with medium-redundant menus than they did with short-redundant menus. The analysis also found a significant three-way interaction effect of configuration × length × item order, with a large effect size ($\eta_p^2$ = 0.204). Analysis of this effect found that the effect of configuration × length is larger when operating inconsistent menus. With the inconsistent condition, the participants were 1.75 s slower in selecting menu targets with medium-complementary menus than they were with short-complementary menus. They were also slower by 1.22 s when operating medium-redundant menus than when operating short-redundant menus. This means that the variance between the two differences (1.75 and 1.22) is 0.53 s in the inconsistent condition. With the consistent condition, the participants were slower in selection time by 0.60 s with the medium-complementary menus compared with short-complementary menus. In addition, they took 0.36 s longer when operating medium-redundant menus than they did when operating short-redundant menus. This means that the variance between the two differences (0.36 and 0.6) is 0.24 s in the consistent condition. The results also demonstrate other significant three-way and four-way interaction effects, but with very small effect sizes (see Table 6).

The tests of between-subjects factors found a significant main effect of item order, and a two-way significant interaction effect of item order × menu format; however, the results did not show a significant main effect of menu format. The main effect of item order means that selection time is significantly different when operating consistent and inconsistent menus. The size of this effect is large ($\eta_p^2$ = 0.653).

**Table 7**
Average post-transitioning selection error and SD.

| Menu format | Menu item order | Menu configuration | Menu length | Average selection error (number of errors per an item selection) | SD |
|---|---|---|---|---|---|
| Text-only | Inconsistent | Complementary | Short | 0.17 | 0.375 |
| | | | Medium | 0.48 | 0.502 |
| | | Redundant | Short | 0.17 | 0.375 |
| | | | Medium | 0.40 | 0.493 |
| | Consistent | Complementary | Short | 0.02 | 0.148 |
| | | | Medium | 0.06 | 0.230 |
| | | Redundant | Short | 0.02 | 0.148 |
| | | | Medium | 0.06 | 0.230 |
| Text-and-icon | Inconsistent | Complementary | Short | 0.11 | 0.316 |
| | | | Medium | 0.36 | 0.481 |
| | | Redundant | Short | 0.11 | 0.316 |
| | | | Medium | 0.32 | 0.470 |
| | Consistent | Complementary | Short | 0.02 | 0.148 |
| | | | Medium | 0.06 | 0.230 |
| | | Redundant | Short | 0.02 | 0.148 |
| | | | Medium | 0.04 | 0.207 |

Participants recorded higher performance in selecting targets with consistent menus. This can be easily observed from the boxplots in Figs. 3 and 4c. Overall, the participants selected menu targets 1.39 s faster when operating consistent menus ($M$ = 1.81 s) than when operating inconsistent menus ($M$ = 3.20 s). The two-way effect of item order × menu format suggests that selection time is influenced by item order and menu format. That is, the effect of inconsistency of item order is larger when users are operating the text-only menus than when they are operating text-and-icon menus. The results showed that participants were 1.65 s faster at selecting targets when operating the consistent menus compared with the inconsistent menus in the text-only condition and were 1.13 s faster at selecting targets with the consistent menus compared with inconsistent menus in the text-and-icon condition. A detailed analysis showed that using text-and-icon decreased the selection time when operating inconsistent menus, but increased the selection time when operating consistent menus. Participants were faster by 0.33 s at selecting targets with text-and-icon inconsistent menus than they were with text-only inconsistent menus. They were also faster by 0.19 s with text-only consistent menus compared with text-and-icon consistent menus. Although the size of item order × menu format effect was moderate ($\eta_p^2 = 0.062$), the results are important and can be considered when designing cross-device menus, particularly menus with inconsistent item order.

Overall, there were a set of variables with large effect sizes such as configuration, length, item order, and length × item order. The results related to the effects of length, item order, and length × item order showed the most practical differences that designers should consider when creating cross-device menus.

### 3.2.2. *Post-transitioning selection error*

We used a mixed ANOVA, with a significance level of 0.05 to analyze error frequency. Table 7 presents the average number of post-transitioning selection errors and standard deviation for all menus tested in the experiment. The overall results showed that users recorded the highest number of errors per item selection with the text-only inconsistent medium-complementary menu ($M$ = 0.48, $SD$ = 0.502) compared with all other tested menus; followed by the text-only inconsistent medium-redundant menu ($M$ = 0.40 s, $SD$ = 0.493). A key result that can be observed from Table 7 is that participants recorded lower selection errors when operating consistent menus than they did when inconsistent menus.

Table 8 shows the main and interaction effects of the independent variables on post-transitioning selection error. Menu configuration had a significant main effect on selection error. The main effect means that selection error was different between the complementary and redundant menus, with a significantly higher selection error found with the complementary menus. The overall average of selection error was 0.16 for complementary menus and 0.14 for redundant menus. However, Table 8 demonstrates that the effect size of configuration is small ($\eta_p^2 = 0.031$), interpreting the small variance between the selection errors for complementary menus and redundant menus.

**Table 8**
Main and interaction effects of menu configuration, menu length, menu item order, and menu format variables on post-transitioning selection error.

| Variable(s) | $F$ | $P$ | $\eta_p^2$ |
|---|---|---|---|
| Configuration | 11.56 | $<0.01$ | 0.031 |
| Length | 69.586 | $<0.001$ | 0.164 |
| Item order | 58.90 | $<0.001$ | 0.142 |
| Menu format | 1.865 | $=0.173$ | 0.005 |
| Configuration × length | 11.56 | $<0.01$ | 0.031 |
| Configuration × item order | 7.74 | $<0.01$ | 0.021 |
| Configuration × menu format | 0.86 | $=0.354$ | 0.002 |
| Length × item order | 42.57 | $<0.001$ | 0.107 |
| Length × menu format | 0.55 | $=0.458$ | 0.002 |
| Item order × menu format | 1.616 | $=0.204$ | 0.005 |
| Configuration × length × item order | 7.74 | $<0.01$ | 0.021 |
| Configuration × length × menu format | 0.860 | $=0.354$ | 0.002 |
| Configuration × item order × menu format | 2.390 | $=0.123$ | 0.007 |
| Length × item order × menu format | 0.334 | $=0.564$ | 0.001 |
| Configuration × length × item order × menu format | 2.390 | $=0.123$ | 0.007 |

Hypothesis df = 1; error df = 356.

The configuration interacted significantly with item order, suggesting that the configuration effect depends on the menu item order, with there being a larger effect of configuration on selection error when operating inconsistent menus. However, the size of this effect is very small ($\eta_p^2 = 0.021$). The overall variance of selection error between complementary menus and redundant menus is 0.03 in the inconsistent condition and 0.01 in the consistent condition. Configuration also interacted significantly with length, demonstrating that the level of the effect of length can depend on the type of menu configuration (complementary versus redundant). This result demonstrates that length has a larger effect on selection error when users operate complementary menus. However, the effect size of configuration × length is also small ($\eta_p^2 = 0.031$). The variance of selection error between medium-complementary menus and short-complementary menus was equal to 0.16, and between medium-redundant menus and short-redundant menus was equal to 0.13.

Length also had a significant main effect on selection error, demonstrating that the number of errors with medium menus was significantly higher than it was with short menus. The average selection error was 0.22 for medium menus and 0.08 for short menus. The interaction of length × item order also had a significant effect on selection error. This interaction effect suggests that the effect of length depends on whether the menu items are ordered consistently across

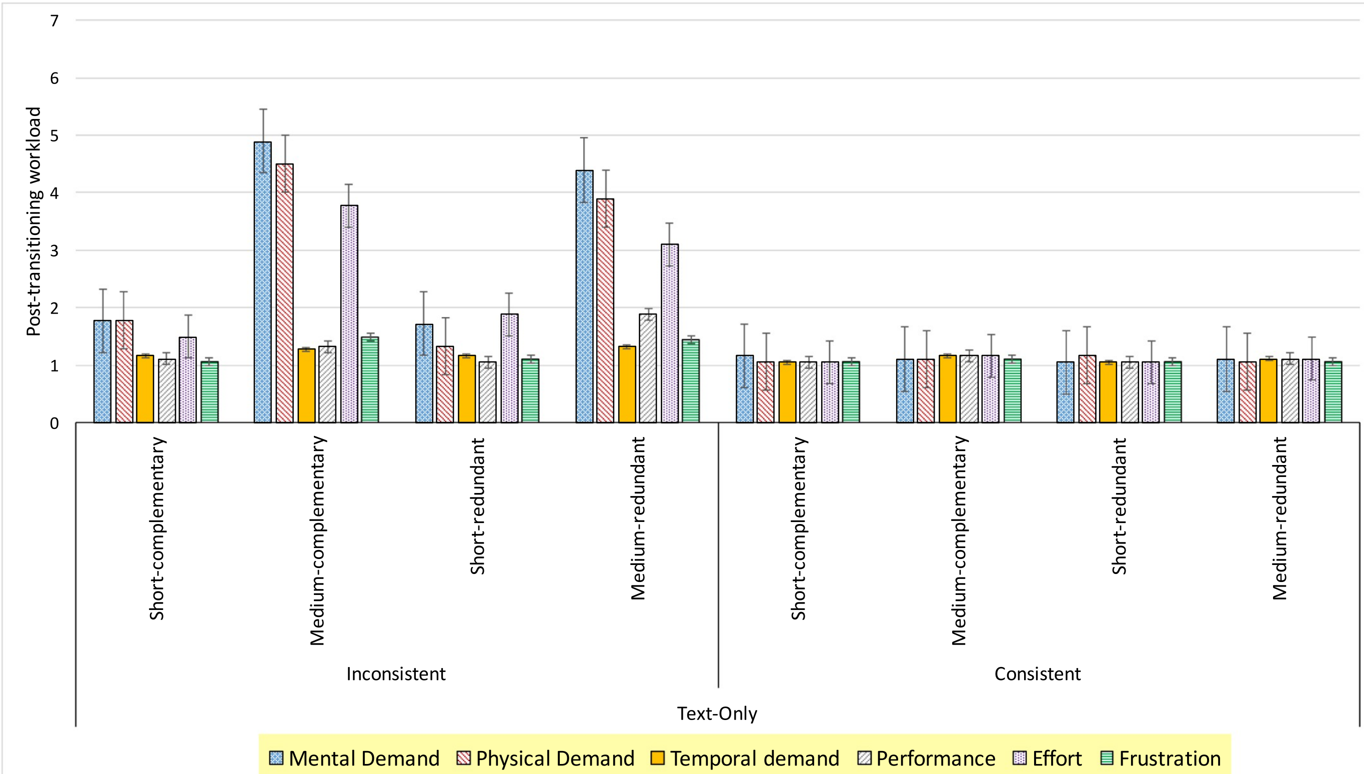



**Fig. 5.** Mean scores of NASA-TLX for post-transitioning workload of the text-only inconsistent and consistent menus.

devices. That is, the selection error for medium menus was much greater than it was with short menus when operating inconsistent menus than when operating consistent menus. The variance of selection error between medium and short menus is 0.25 for the inconsistent condition and 0.035 for the consistent condition. An important result from Table 8 is that there is a large effect of length ($\eta_p^2 = 0.164$) and an almost-large effect of length × item order ($\eta_p^2 = 0.107$) on selection error.

The statistics also found a three-way interaction: configuration × length × item order, demonstrating that the effect of configuration × length depends on the consistency of item order, with a greater effect present when operating inconsistent menus. However, the size of the interaction effect was small ($\eta_p^2 = 0.021$).

The analysis of the main and interaction effect of the between-subjects factors (menu item order and menu format) demonstrated a significant main effect only for menu item order. The size of this effect was large ($\eta_p^2 = 0.142$). This effect means that the selection error was much lower when participants operated consistent menus. The overall average of selection errors for consistent menus was 0.04 and for inconsistent menus was 0.26.

In general, the variables with large effect sizes were menu length and menu item order. The results related to these variables showed the most practical differences that designers should consider when creating cross-device menus. The finding for the interaction effect of length × item order also showed a practical difference that should also be considered by designers.

### 3.2.3. NASA-TLX

We analyzed the scores of the six indices of the NASA-TLX. We used the NASA-TLX to assess users' post-transitioning workload, that is, the level of task load users reported when they had switched to another device to re-locate menu targets. A general observation is that the workload, in particular the five indices of the NASA-TLX (mental demand, physical demand, performance, effort, and frustration) were higher with medium inconsistent menus (medium-complementary and medium-redundant) across the different menu formats: text-only (Fig. 5) and text-and-icon (Fig. 6).

We used one-way ANOVA, and the Scheffe post hoc test to compare the workload between short and medium menus within each experimental condition, and found significant differences among inconsistent menus (see Table 9). The level of mental demand, physical demand, and effort for medium-complementary menu and medium-redundant menu was significantly higher than it was with the short-complementary menu and the short-redundant menu. These results were observed across the two menu formats (text-only and text-and-icon). The levels of performance and frustration were also significantly higher with medium menus (medium-complementary or medium-redundant) compared with short menus (short-complementary and or short-redundant). The temporal demand was not significantly different between the short and medium inconsistent menus. With the consistent menus across the different menu formats, there were no significant differences in the six indices of the NASA-TLX between the medium and short menus.

A one-way ANOVA and Scheffe post hoc test were also conducted to compare the workload of each menu (short-complementary, medium-complementary, short-redundant menu, and medium-redundant) across the different menu designs (text-only inconsistent menus, text-only consistent menus, text-and-icon inconsistent menus, and text-and-icon inconsistent menus). For the short-complementary menus, we found significant results for three indices: mental demand, physical demand, and effort (see Table 10). Analysis of these results demonstrated that the level of mental demand, physical demand, and effort when operating inconsistent short-complementary menus across the two different menu formats (text-only and text-and-icon) was significantly higher than it was when operating consistent short-complementary menus.

The multiple-comparison test of short-redundant menus found significant results for the level of mental demand, and effort (see Table 10). These findings indicate that the level of mental demand and effort with text-only inconsistent short-redundant menus is significantly

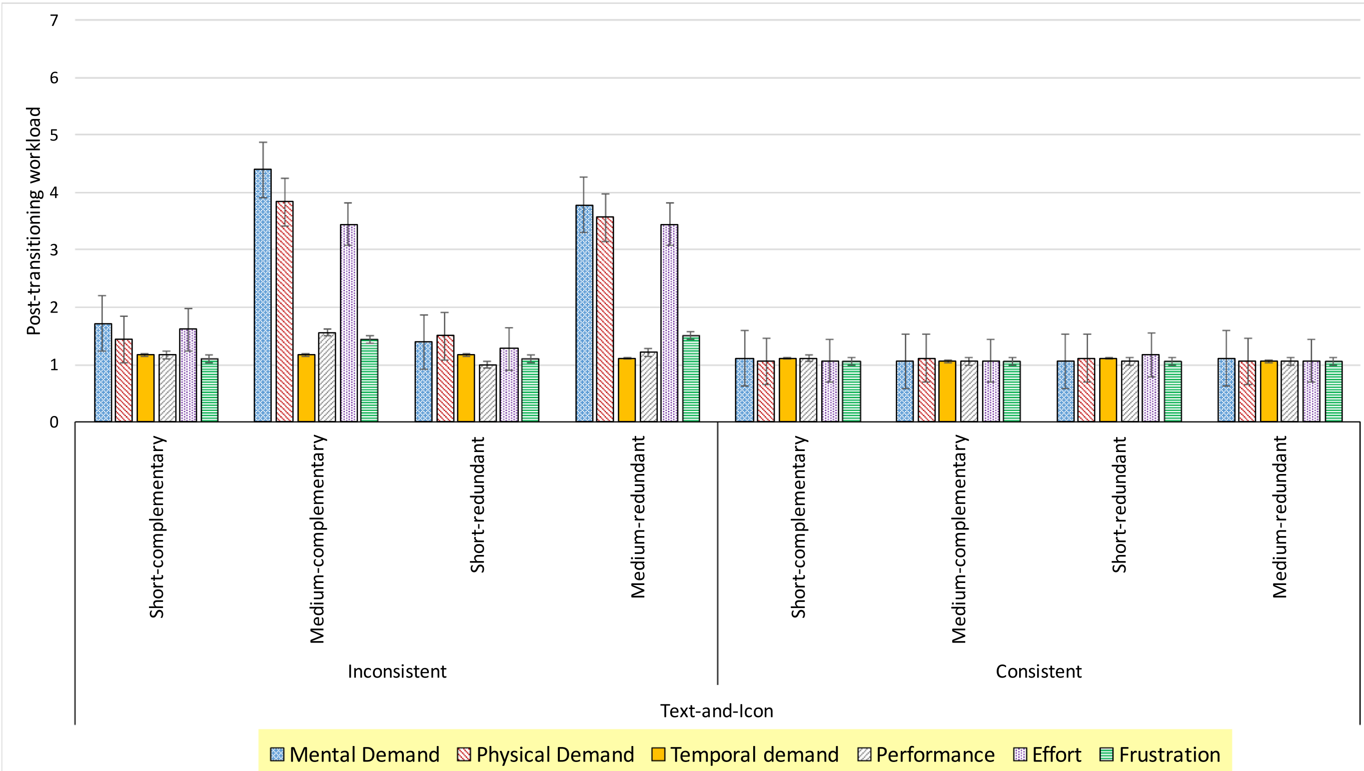



**Fig. 6.** Mean scores of NASA-TLX for post-transitioning workload of the text-and-icon inconsistent and consistent menus.

higher than it is with text-only and text-and-icon consistent short-redundant menus.

For the medium-complementary menus, we found significant results for five indices of the NASA-TLX: mental demand, physical demand, performance, effort, and frustration (see Table 10). These robust significant results demonstrate that the level of mental demand, physical demand, and effort when operating text-only and text-and-icon inconsistent medium-complementary menus is significantly higher than it is when operating text-only and text-and-icon consistent medium-complementary menus. The level of performance when operating a text-and-icon inconsistent menu was significantly higher than it was when operating a text-and-icon consistent menu. The level of frustration when operating a text-only inconsistent menu was significantly higher than it was when operating a text-and-icon consistent menu.

For the medium-redundant menus, we also found significant results for the NASA-TLX indices of mental demand, physical demand, performance, effort, and frustration (see Table 10). The level of mental demand, physical demand, effort, and frustration was significantly higher when operating inconsistent medium-redundant menus than it was when operating consistent medium-redundant menus. The level of performance when operating text-only inconsistent medium-redundant menu was significantly higher than it was when operating other medium-redundant menus (text-and-icon inconsistent medium-redundant, text-only consistent medium-redundant, and text-and-icon consistent medium-redundant).

The main interpretation of results is that menu item order (inconsistent versus consistent) significantly affects the results of most of the indices of the NASA-TLX. That is, inconsistent menus were found to require greater workload from users than consistent menus. Another key finding is that medium menus required greater workload from participants than short menus.

**Table 9**
One-way ANOVA and Scheffe post hoc test results of NASA-TLX for post-transitioning workload of short and medium menus within each experimental condition.

| | Text-only | | | | Text-and-icon | | | |
|---|---|---|---|---|---|---|---|---|
| | Inconsistent | | Consistent | | Inconsistent | | Consistent | |
| | *F* | Scheffe post hoc | *F* | Scheffe post hoc | *F* | Scheffe post hoc | *F* | Scheffe post hoc |
| Mental demand | $33.632^{***}$ | MC, MR > SC, SR | $0.360^{n.s.}$ | – | $26.717^{***}$ | MC, MR > SC, SR | $0.231^{n.s.}$ | – |
| Physical demand | $35.469^{***}$ | MC, MR > SC, SR | $0.562^{n.s.}$ | – | $22.597^{***}$ | MC, MR > SC, SR | $0.231^{n.s.}$ | – |
| Temporal demand | $0.674^{n.s}$ | – | $0.562^{n.s.}$ | – | $0.102^{n.s.}$ | – | $0.231^{n.s.}$ | – |
| Performance | $9.581^{***}$ | MR > SC, SR | $0.562^{n.s.}$ | – | $22.597^{**}$ | MC > SR | $0.205^{n.s.}$ | – |
| Effort | $21.039^{***}$ | MC, MR > SC, SR | $0.562^{n.s.}$ | – | $31.352^{***}$ | MC, MR > SC, SR | $0.705^{n.s.}$ | – |
| Frustration | $5.397^{**}$ | MC > SC | $0.205^{n.s.}$ | – | $3.263^{*}$ | MR > SC, SR | $0.000^{n.s.}$ | – |

SC = short-complementary, MC = medium-complementary, SR = short-redundant, MR = medium-redundant.
$^{n.s}$ > 0.05.
$^{*}$ $p < 0.05$.
$^{**}$ $p < 0.01$.
$^{***}$ $p < 0.001$.

**Table 10**
One-way ANOVA test results of NASA-TLX for post-transitioning workload of each menu across the different menu designs.

| | Short-complementary | Short-redundant | Medium-complementary | Medium-redundant |
|---|---|---|---|---|
| | *F* | *F* | *F* | *F* |
| Mental demand | 13.966*** | 12.750*** | 52.740*** | 37.965*** |
| Physical demand | 11.052*** | 2.556[n.s.] | 45.057*** | 41.705*** |
| Temporal demand | 0.448[n.s.] | 0.731[n.s.] | 0.576[n.s.] | 2.185[n.s.] |
| Performance | 0.360[n.s.] | 0.333[n.s.] | 2.912* | 11.676*** |
| Effort | 5.599** | 9.847*** | 71.207*** | 34.912*** |
| Frustration | 0.205[n.s.] | 0.231[n.s.] | 4.019* | 6.597** |

[n.s] > 0.05.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

### 3.2.4. Learnability

As stated, the participants operated each menu three times. We analyzed the learning curve for all menus based on the initial and the repeated operations. Fig. 7 presents the learning curve of all the types of menus: (a) text-only inconsistent; (b) text-only consistent; (c) text-and-icon inconsistent; (d) text-and-icon consistent. Given that our study focuses on post-transitioning user performance, the results in Fig. 7 are based on the selection time of users after transitioning from one device to another. Overall, Fig. 7 demonstrates that the post-transitioning user selection times show a decreasing trend from the first to the second menu operations for all menus.

For the text-only inconsistent menus, both the short and medium menus showed a slight decrease in the selection times in the second operation; however, with the third operation, the decrease does not continue, except when using the short-complementary menu, where a very small decrease was found (−0.04 s). The variance from the first to the third operation was −0.24 s for short-complementary menu, −0.36 s for medium-complementary menu, −0.18 s for short-redundant menu, and −0.25 s for medium-redundant menu. Although the decrease was higher for the medium menus, the decreases did not reach or even come close to the average selection times for the short menus in the third operation.

For the text-and-icon inconsistent menus, there was a continuous decreasing trend from the first to the third operation for all menus. The variance from the first to the third operation was −0.32 s for the short-complementary menu, −0.58 s for the medium-complementary menu, −0.49 s for the short-redundant menu, and −0.59 s for the medium-redundant menu. Compared with the text-only inconsistent menus, the variance from the first to the third operation was greater for the text-and-icon menus. This means that using text-and-icon menu items helped users to learn inconsistent menus more than did using text-only menu items.

The variable of configuration seemed to continue to play a role in the efficiency of menu-target selection in the repeated operations of menus with inconsistent item order. When comparing text-only short menus (short-complementary versus short-redundant), and text-only medium menus (medium-complementary versus medium-redundant), the average selection times for inconsistent complementary menus in the third operation were higher than the average selection times for the redundant menus. The average selection time in the third operation was 2.57 s for the short-complementary menu, 2.33 s for the short-redundant menu, 4.13 s for the medium-complementary menu, and 3.42 s for the medium-redundant menu. Two further observations are that for text-and-icon inconsistent menus: (1) the variance between the first and third operation for short menus is smaller with the short-complementary menu (−0.32 s) than it is with the short-redundant menu (−0.49 s); (2) the average selection time for the medium-complementary menu (3.40 s) was higher than it was with the medium-redundant menu (3.07 s).

With text-only consistent menus, the selection times of medium-length menus (medium-complementary, and medium-redundant) were higher than they were with the short menus (short-complementary and short-redundant) in the first operation. However, the average selection times of the short and medium menus came very close to each other in the second and third operations, with averages in the third operation of 1.12 s for the short-complementary menu; 1.22 s for the medium-complementary menu; 1.01 s for the short-redundant menu; and 1.18 s for the medium-redundant menu. The same pattern can be observed with text-and-icon consistent menus, where the short and medium menus had much closer average selection times in the third operation: 1.06 s for the short-complementary menu, 1.30 s for the medium-complementary menu, 1.03 s for the short-redundant menu, and 1.13 s for the medium-redundant menu. Although the complementary menus had marginally higher average selection times than did the redundant menus and the medium menus had slightly higher average selection times than did the short menus, the differences in learning efficiency with the consistent menus were not very great when compared with the differences in learning efficiency with the inconsistent menus.

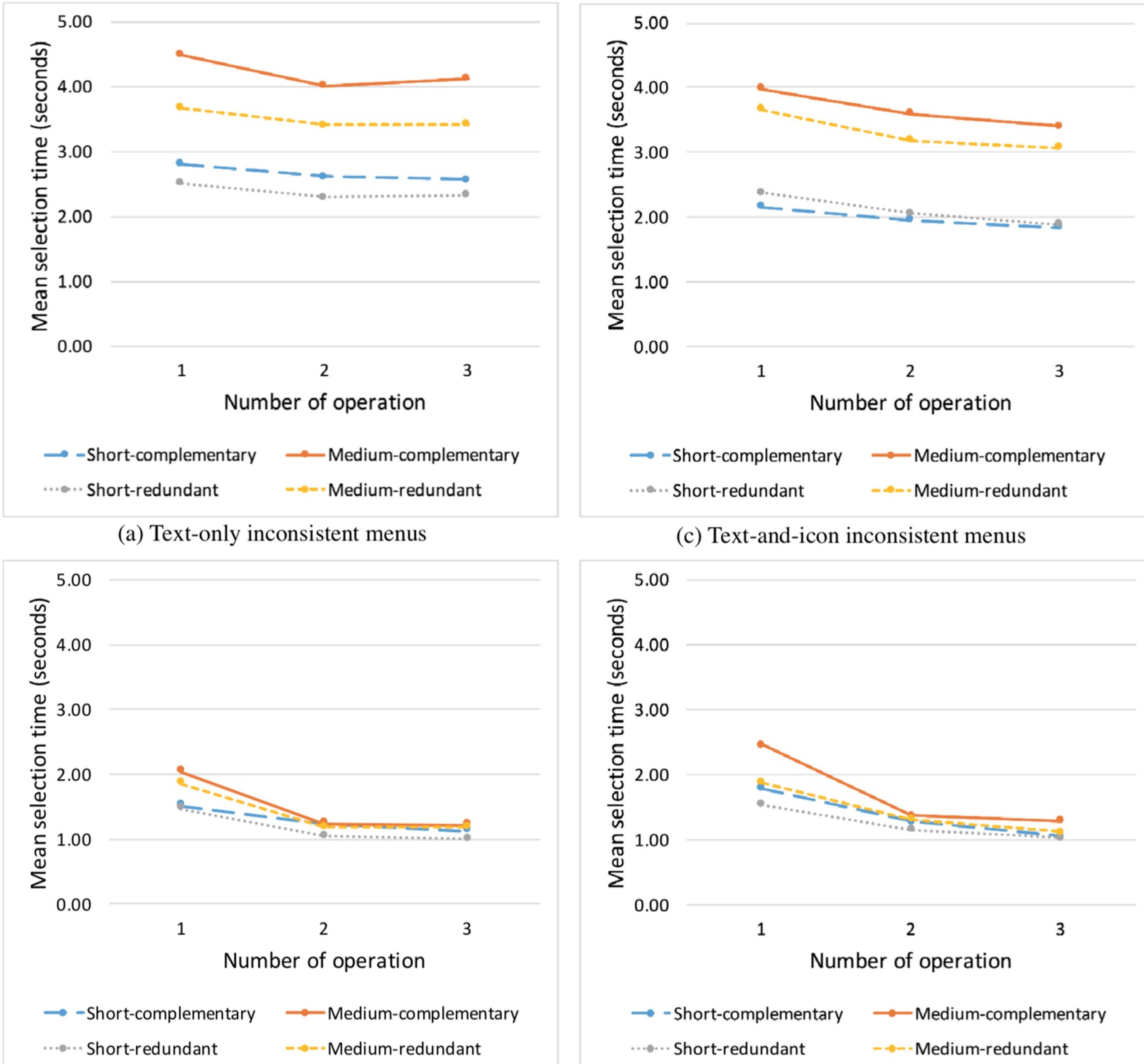
## 4. Experiment 2

We further investigated the menu interfaces of the 46 cross-platform services from our informal analysis conducted before Experiment 1. We found that 77% of all analyzed menus were organized horizontally for the desktop interfaces and vertically for the mobile interfaces, and that 23% of the menus applied a vertical layout for the desktop and mobile interfaces. Thus, we designed another experiment to compare post-transitioning user performance on cross-device menu interfaces using horizontal–vertical layout and vertical–vertical layout at two different lengths (short and medium). Therefore, we designed eight desktop and mobile menu interfaces, representing four different designs of cross-device menu interfaces: short horizontal–vertical menu; medium horizontal–vertical menu; short vertical–vertical menu; and medium vertical-vertical menu. Table 11 presents the independent variables and their levels, and the menus used in Experiment 2.

In this section, we present the experimental menu-interface design, the participants and experimental design, and the experimental results for Experiment 2. For this experiment, we had the same apparatus, measures and stimulus, and procedures used in Experiment 1. However, NASA-TLX was not used in Experiment 2 due to the limited time allocated for the sessions.

### 4.1. Menu design

For Experiment 2, we used the text-only consistent redundant menu interfaces developed in Experiment 1 with the two different lengths (short and medium). These menus were designed with a horizontal (desktop) vertical (mobile) layout (see Fig. 1). We also designed two vertical menus with the two different lengths (short and medium) for the desktop interfaces. For these vertical menus, the vertical separation




(a) Text-only inconsistent menus

(c) Text-and-icon inconsistent menus

(b) Text-only consistent menus

(d) Text-and-icon consistent menus

**Fig. 7.** Learning curve of participants for all menus in Experiment 1.

between items was approximately 20 pixels, as used in previous studies (e.g., Hornof and Kieras, 1997). The first menu item was 80 pixels from the top of the screen and the menu items were 200 pixels from the right of the screen. Fig. 8 presents an example of the English version of the short vertical–vertical menu for the desktop and mobile interfaces.

### 4.2. Participants and experimental design

We had 20 undergraduate students (10 females and 10 males), aged 20 to 24 years with over two years of internet and cross-device experience. All participants had high-school certification as their highest education level. All participants, who joined Experiment 2, did not take part in Experiment 1. Participants did not receive any form of compensation for taking part in this experiment. We used a within-subjects design to compare the two redundant menu layouts: horizontal–vertical and vertical–vertical with two different menu lengths (short and medium). To reduce the effects of order of menus, the participants were divided into four groups (A, B, C, and D) of five participants. Each group searched all four cross-device menus: short horizontal–vertical, medium horizontal–vertical, short vertical–vertical, and medium vertical–vertical (see Table 12).

Each group operated menus in a different order. Each participant operated the menus in two sessions (one session for short menus and the other for medium menus). The order of the menu layout was counterbalanced within each session. The order of menu length was also counterbalanced within the sessions. The items were ordered differently within each menu layout of each length (short and medium), but the positions of the targeted items across the two conditions (horizontal–vertical and vertical–vertical) were controlled. Each participant conducted 48 cross-device search trials, divided into four blocks (the

**Table 11**
Independent variables and their levels, and the menus in Experiment 2.

| | | Menu length<br>Short | Medium |
|---|---|---|---|
| Menu layout | Horizontal–vertical | Short horizontal–vertical menu | Medium horizontal–vertical menu |
| | Vertical–vertical | Short vertical–vertical menu | Medium vertical–vertical menu |

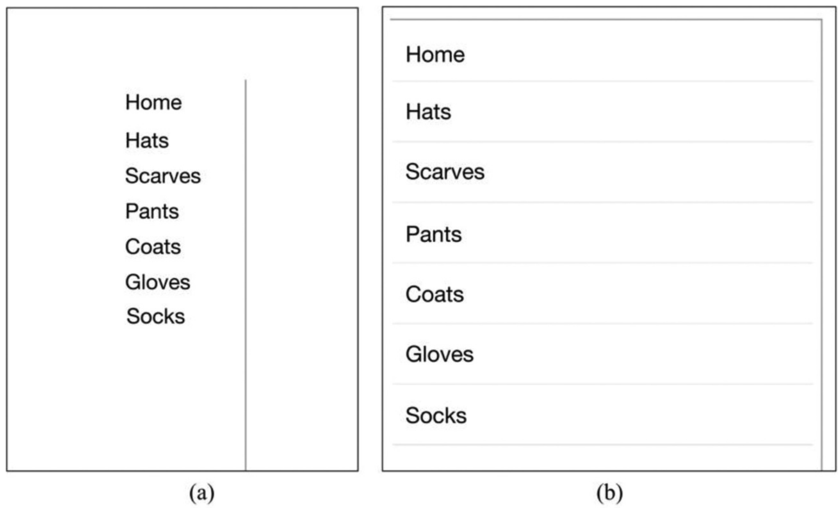



**Fig. 8.** Illustration of the English version of the short vertical-vertical menu for the desktop (a) and mobile (b) interfaces.

**Table 12**
Design of Experiment 2.

| Group | Session 1 | | Session 2 | |
|---|---|---|---|---|
| A | Short horizontal–vertical | Short vertical-vertical | Medium horizontal–vertical | Medium vertical-vertical |
| B | Short vertical-vertical | Short horizontal–vertical | Medium vertical-vertical | Medium horizontal–vertical |
| C | Medium horizontal–vertical | Medium vertical-vertical | Short horizontal–vertical | Short vertical-vertical |
| D | Medium vertical-vertical | Medium horizontal–vertical | Short vertical-vertical | Short horizontal–vertical |

number of tasks in each block depended on the menu layout and length). Each cross-device search trial attempted on each menu involved a device change, resulting in two tasks per trial (see Section 3.1.5 for more detail). Participants operated each menu three times to investigate the learning-curve efficiency of operating the menus. Thus, we had 20 participants × 48 trials × 2 tasks per trial × 3 operations of each menu = 5760 selections. Half of these selections (2880) were associated with re-location of menu items after switching from one device to another. Our analysis of post-transitioning user performance was based on these selections associated with re-location of menu items. The entire experiment required approximately 30 min to complete.

### *4.3. Analysis and results*

In this section, we analyze our experimental results of the effects of menu layout (horizontal-vertical and vertical-vertical) and menu length (short and medium) on post-transitioning selection time and selection error of menu items. We also analyze the results of the learning curve of the menus.

To analyze the results, a repeated measure ANOVA was used, with a significance level of 0.05. Table 13 presents the average post-transitioning selection times and selection error for the short and medium menus across the two layouts (horizontal–vertical and vertical–vertical). The overall results demonstrated that the average selection times for the short menus across the two layouts were similar. In addition, participants recorded only 0.10 s faster selection time with the medium vertical–vertical menu than with the medium horizontal–vertical menu. Fig. 9 illustrates that the median of the selection time of the medium horizontal–vertical menu is slightly higher than the median of the selection time of the short horizontal–vertical menu. However, in the vertical–vertical condition, the median values for the short and medium menus are very similar.

Table 14 shows the main and interaction effects of the independent variables on post-transitioning selection time and selection error. We found that there was no significant effect of layout on selection time, meaning that selection time for locating the menu targets post-transition was not significantly different between horizontal–vertical and vertical–vertical menus. This can be easily observed in Fig. 10a. There was a significant main effect of length on selection time, demonstrating that users are slower at re-locating targets with medium menus. However, the size of this effect was very small ($\eta_p{}^2 = 0.016$). That is, users were slower by only 0.08 s when operating the medium menus. Fig. 10b demonstrates the similarity of the median values for the short and medium menus.

The variable of menu layout interacted significantly with the variable of menu length, suggesting that the effect of length depends on whether the menu layout is different across devices. This means that the effect of length is larger when operating horizontal–vertical menus when the layout is inconsistent across devices. However, the effect size was small (ηp2 = 0.035). The difference between the average selection

**Table 13**
Average post-transitioning selection time and error.

| Layout | Menu length | Average selection time (time taken in seconds per an item selection) | SD | Average selection error (number of errors per an item selection) | SD |
|---|---|---|---|---|---|
| Horizontal–vertical | Short | 1.37 | 0.497 | 0.01 | 0.177 |
| | Medium | 1.50 | 0.561 | 0.03 | 0.165 |
| Vertical-vertical | Short | 1.38 | 0.439 | 0.04 | 0.201 |
| | Medium | 1.40 | 0.471 | 0.04 | 0.201 |

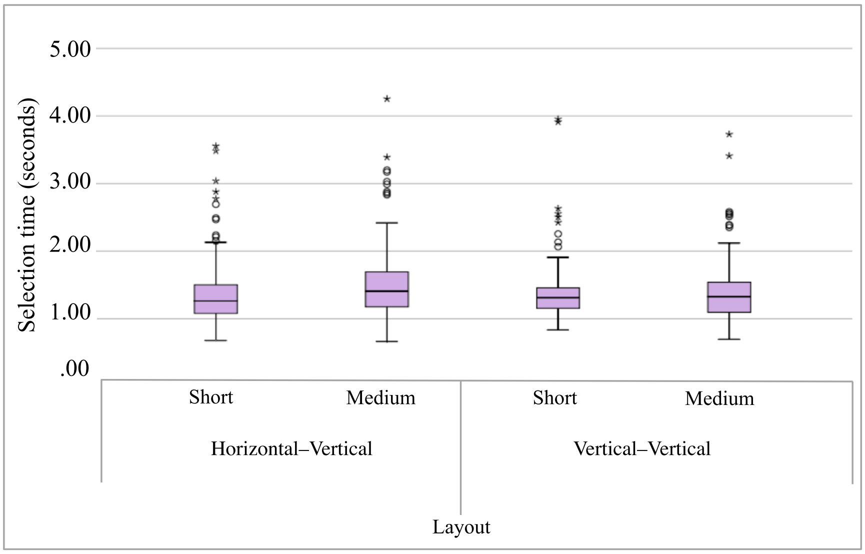



**Fig. 9.** Boxplot of post-transitioning selection time in seconds for all menus in Experiment 2.

time for short horizontal–vertical menu and medium horizontal–vertical menu was only higher by 0.11 s than the difference between the average selection time for short vertical–vertical menu and medium vertical–vertical menu (see Fig. 9).

The post-transitioning selection error was generally very low for each menu and the accuracy was always more than 96% (see Table 13). We found that there was no significant main or interaction effect of layout and length on selection error (see Table 14).

Fig. 11 presents the learning curve of the short and medium menus across the two layouts (horizontal–vertical and vertical–vertical). There was a decreasing trend in the average selection time in the second and third operations for all menus. The average selection time for medium horizontal–vertical menu was higher than the average selection time of all other menus in the first operation. In the second and third operations, the average selection time for the medium horizontal–vertical menu became very close to the average selection time for the other menus. The variance between the first and third operation for the medium horizontal–vertical menu was −0.29 s. The average selection time in the third operation was 1.15 s for the short horizontal–vertical menu, 1.21 s for the medium horizontal–vertical menu, 1.14 s for the short vertical-vertical menu, and 1.17 s for the medium vertical-vertical menu.

## 5. Discussion

The principal aim of this study was to examine user performance on cross-device menu interfaces and identify the effects of a set of variables on post-transitioning selection efficiency and selection error. The study also investigated the learnability and NASA-TLX of the cross-device menu interfaces. Table 15 summarizes the main and interaction effects of the variables on post-transitioning selection time and selection error. In Table 15, variables with large effect sizes ($\eta_p^2 \geq 0.14$) are highlighted in bold. In this section, we focus the discussion on the main and interaction effects with large effect sizes. We also highlight the theoretical and practical implications of the study and potential future research.

### 5.1. Menu configuration

The variable of menu configuration significantly affects selection time and selection error. Users were generally faster at selecting menu items on redundant menus than on complementary menus. The selection error was also higher for complementary menus. However, the effect size of menu configuration on selection error was not large. Repeated operations with the experimental menu interfaces designed

**Table 14**
Main and interaction effects of menu layout, and menu length variables on post-transitioning selection time and selection error.

| | Post-transitioning selection time | | | Post-transitioning selection error | | |
|---|---|---|---|---|---|---|
| Variable(s) | *F* | *P* | $\eta_p^2$ | *F* | *P* | $\eta_p^2$ |
| Layout | 2.26 | =0.134 | 0.016 | 3.04 | =0.083 | 0.021 |
| Length | 7.27 | <0.01 | 0.048 | 3.04 | =0.083 | 0.021 |
| Layout × length | 5.71 | <0.05 | 0.035 | 0.33 | =0.566 | 0.002 |

Hypothesis df = 1; error df = 143.

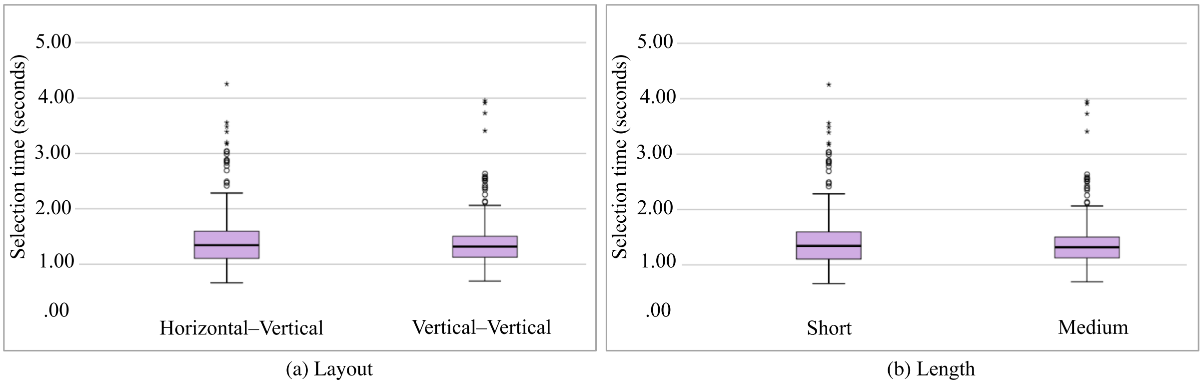



**Fig. 10.** Boxplot of post-transitioning selection time in seconds for the levels in each independent variable in Experiment 2.

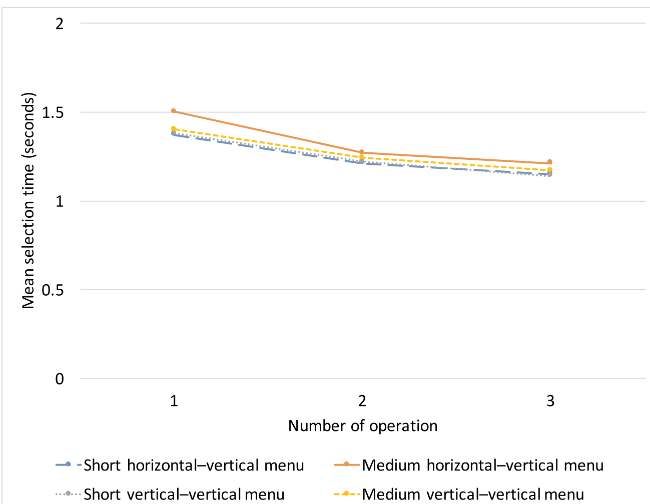



**Fig. 11.** Learning curve of participants for all menus in Experiment 2.

with inconsistent item order were found to have a lower level of learning efficiency with complementary menus than they did with redundant menus. The effect of the complementary configuration was because of the menu items not being fully redundant across the devices. This means that users might find it difficult to cope with cross-device menus that have a mixture of shared and unique items, particularly when the shared items are not ordered consistently across the devices. Therefore, it can be advised that it is better to design cross-device redundant menus.

### *5.2. Menu length and configuration*

While prior studies (e.g., Byrne et al., 1999; Halverson and Hornof, 2008; Nilsen, 1991) have identified the effect of menu length when operating single menu interfaces, our study revealed that length can also significantly affect post-transitioning user performance when operating multiple menu interfaces across devices. Participants were faster at selecting items and had a higher level of selection error with medium menus compared with short menus. The length had the largest effect size among all the variables on both selection time and selection error. We also found that length interacted significantly with configuration, demonstrating a larger effect of length on selection time when users operate complementary menus compared with when they operate redundant menus. These results mean that designers should consider that it is better to design cross-device short menus than cross-device medium menus. However, when it is necessary to design cross-device medium menus, it is recommended to configure the menus with redundant menu items.

**Table 15**
Summary of main and interaction effects of the independent variables on post-transitioning selection time and error.

| | Variable(s) | Post-transitioning selection time | $\eta_p^2$ | Post-transitioning selection error | $\eta_p^2$ |
|---|---|---|---|---|---|
| Exp. 1 | **Configuration** | *** | **0.632** | ** | 0.031 |
| | **Length** | *** | **0.972** | *** | **0.164** |
| | **Item order** | *** | **0.653** | *** | **0.142** |
| | Menu format | n.s. | 0.004 | n.s. | 0.005 |
| | **Configuration × length** | *** | **0.623** | ** | 0.031 |
| | Configuration × item order | n.s. | 0.009 | ** | 0.021 |
| | Configuration × menu format | *** | 0.050 | n.s. | 0.002 |
| | **Length × item order** | *** | **0.900** | *** | 0.107 |
| | Length × menu format | *** | 0.066 | n.s. | 0.002 |
| | Item order × menu format | *** | 0.062 | n.s. | 0.005 |
| | **Configuration × length × item order** | *** | **0.204** | ** | 0.021 |
| | Configuration × length × menu format | *** | 0.034 | n.s. | 0.002 |
| | **Configuration × item order × menu format** | *** | **0.475** | n.s. | 0.007 |
| | Length × item order × menu format | ** | 0.021 | n.s. | 0.001 |
| | Configuration × length × item order × menu format | ** | 0.021 | n.s. | 0.007 |
| Exp. 2 | Layout | n.s. | 0.016 | n.s. | 0.021 |
| | Length | ** | 0.048 | n.s. | 0.021 |
| | Layout × menu length | * | 0.035 | n.s. | 0.002 |

n.s $> 0.05$.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

### 5.3. Menu item order, configuration, and length

Our study demonstrated that users perform better when operating menus that use consistent item order across devices and that the effect of length and configuration × length can significantly decrease when users operate consistent menus compared with when they operate inconsistent menus. While spatial consistency is an aspect of menu design that can help predict total selection time in the context of interaction with a single menu within an individual device (Bailly et al., 2014), our results highlight the importance of the spatial consistency of menu items across devices for better post-transitioning user performance. The results also confirm Majrashi's (2016) previous finding that users rely greatly on spatial memory when re-locating user-interface elements when transitioning between user interfaces across devices. The results of the NASA-TLX also agree with the results for user performance in our study, suggesting that the workload for re-locating items on inconsistent menus is significantly greater than it is for re-locating items on consistent menus. In addition, the results suggest that the workload indices mental demand, physical demand, and effort are significantly greater with medium menus than they are with short menus when the item order is inconsistent. However, when the item order is consistent, there is no significant difference between mental demand, physical demand, and effort at different menu lengths. The results also indicate that menus with inconsistent item order have a less efficient learning curve than menus with consistent item order. The selection error was also significantly lower with the consistent menus than it was with the inconsistent menus. This could be because of the habituation behavior identified in Majrashi (2016). That is, some user actions applied on a user interface of a specific service on a specific device become reflexive when using an interface of the service on another device. According to Majrashi (2016), a user can reflexively select an item within a specific list after transitioning to another interface without reading the label of the item, relying on the position of the item as experienced with the previous interface. This can occur particularly when the user switches devices in a short time.

In general, the conclusions we draw here is that if the order of menu items is consistent, the configuration and length do not have a large effect on post-transitioning user performance, workload, and learnability. Therefore, designers can use any combination of configuration (redundant or complementary) and length (short or medium) when designing cross-device menus as long as the order of menu items is consistent. Here, it is important to note that while complementary menus involve shared and non-shared menu items across devices, the positioning of shared items should start from the top of menus, followed by the non-shared items at the bottom of the menus. In addition, the shared items should have a consistent order of items across devices.

### 5.4. Menu format, item order, and configuration

Neither post-transitioning selection time nor selection error was significantly affected by changing the menu format. However, there was a significant two-way interaction effect of menu item order and menu format, with a moderate effect size. Participants were faster at selecting menu targets with text-and-icon inconsistent menus than they were with text-only inconsistent menus. They also were faster with text-only consistent menus compared with text-and-icon consistent menus. In addition, an interaction effect of configuration × item order × menu format was identified with a large effect size, demonstrating a higher selection time for complementary menus than for redundant menus when users operated text-only inconsistent menus compared with text-and-icon inconsistent menus, and when they operated text-and-icon consistent menus compared with text-only consistent menus. These results indicate that users might employ visual memory to re-locate items when operating inconsistent menus given that spatial memory is not suitable for re-locating items when using inconsistent menus. This interpretation is supported by Majrashi (2016), who found that users can employ visual memory to re-find elements on another device "by remembering what they looked like, rather than their names." Majrashi (2016) concludes with a design recommendation that if user-interface elements cannot be placed in the same positions across devices, the elements should at least have a consistent visual appearance. This conclusion could mean that the users in our study employed visual memory rather than spatial memory when elements were placed inconsistently across devices. As a consequence, users had a faster selection time and a smaller effect of configuration with text-and-icon inconsistent menus compared with text-only inconsistent menus. However, when operating consistent menus, users might use spatial memory for re-locating targets and that the use of visual items (icons) seems to increase the time taken to select the menu targets. The selection error was also statistically lower with the text-and-icon inconsistent menus than it was with text-only inconsistent menus, suggesting that using icons in inconsistent menus can reduce the error rate, confirming a finding of an early study by Kacmar and Carey (1991). The use of icons with text in menus can also enhance the learning efficiency of inconsistent menus compared with using text-only menus, as found in our study.

In general, these results can inform the designers that there is no need to use icons beside texts for menu items when the menus apply a consistent order of items across devices. However, when designers need to construct menus with an inconsistent order of items across devices, icons should be used in combination with text for the menu items.

### 5.5. Menu layout and length

In Experiment 2, the variable menu layout did not have a significant effect on user performance; however, there was a significant effect of menu length, confirming the results for the effect of menu length in Experiment 1. However, the effect size of menu length in Experiment 2 was not as large as it was in Experiment 1, which might be because all the tested menus in Experiment 2 had a consistent order of items. This interpretation is supported by our finding related to the interaction effect of length and item order in Experiment 1, where length had a smaller effect when users operated consistent menus compared with inconsistent menus. Length also interacted significantly with layout, suggesting a larger effect of length when the layout of the menu was different across devices (a horizontal layout on the desktop interface and a vertical layout on the smartphone interface); however, the effect size was small. The learnability results in Experiment 2 showed a similar pattern of learning curve for all menus. This pattern was similar to the learning curve of the menus that had a consistent item order in Experiment 1. Such similarity is because all menus investigated in Experiment 2 applied a consistent item order.

In general, all the effect sizes for Experiment 2 were small. Consequently, we conclude that if the order of menu items is consistent, then the layout and length do not have a large effect on post-transitioning user performance and learnability. This result is important because it informs designers that they can use any combination of layout (horizontal–vertical or vertical-vertical) and length (short or medium) of menus across devices as long as the menus apply a consistent order of items.

### 5.6. Theoretical and practical implications

Our study has theoretical and practical implications. Theoretically, the study explains the observed effects of variables on post-transitioning user performance, and discusses possible related behaviors influencing re-locating menu items on another device. It also extends the current research on user performance with menu interfaces from the context of one-menu-one-device to the context of multi-menu-multi-device. Our analysis and identification of the different types of cross-device menu interfaces (redundant and complementary) and different independent variables (e.g., menu length, menu item order and label consistency, menu format, and menu layout), as well as our study results provide the basis and insights for future research on multi-device menu interaction.

Task continuity and seamless transitioning between devices are important elements of the cross-device user experience and inter-usability (Denis and Karsenty, 2004; Wäljas et al., 2010). Theoretically, our results demonstrate that the design of menus across devices is important in supporting task continuity and seamless transitioning between devices. That is, menus that allow more efficient re-location of items after the transition from one device to another can support continuity and seamless transitioning. Further, our findings in Experiment 1 in relation to the importance of consistency for menu item order shed light on the debate highlighted by Majrashi (2016) surrounding the importance of consistency in cross-device user interaction. In Experiment 2, we found that the consistency of menu layouts across devices had no effect on user performance. Thus, our findings add value to the ongoing debate about the importance of consistency by associating it with real-world user needs and behaviors. For example, consistency of menu item order is important for cross-device design because it helps improve user performance, and the users in our study seemed to be employing spatial memory as a behavior when searching menu items across devices. However, the consistency of menu layout across the devices was less important. That is, the cross-device menus can have consistent or inconsistent layout as long as the menus apply a consistent order of items.

From the practical perspective, designers can use our study results and the conclusions we draw here as a resource for designing cross-device menus that allow more efficient and accurate selection of menu items after the transition from one device to another. The designers can also rely on our results to ensure reducing workload and improving learnability of cross-device menus. For example, users had a better performance, a more efficient learning curve, and a lower NASA-TLX workload with cross-device menus that have consistent item order. Therefore, designers should consider the importance of consistency of item order when designing menu interfaces of multi-device interactive systems that can be targeted for use in cross-device interaction contexts. In addition, the different menu designs tested in our study can provide designers with better options when they are restricted to specific menu designs that are driven by the design and content of interfaces across devices. For example, when the content of the system (data and functions) is not fully accessible on all devices, and each device has different specific content priorities, designers may need to construct complementary menus with an inconsistent item order. According to our findings, the most appropriate option for a menu design in this case would be to use a combination of text and icons for menu items across devices.

For the purpose of our study, we used the averaged post-transitioning selection time and selection error per single menu-target selection when comparing the user performance across the tested menus. Therefore, the variances between the selection times and between the selection errors of the different menus could be considered small, and it could be argued that they reveal no practical differences between menu designs. However, calculating the overall menu operating time and operating error shows a clear difference between user performances on the different types of menus. For example, users in Experiment 1 spent a total of 2313.29 s (38.55 min) on selecting menu targets on menus with an inconsistent order of items compared with 1312.47 s (21.87 min) on selecting menu targets on menus with a consistent order of items. In addition, the total number of selection errors was 190 for inconsistent menus compared with only 27 for consistent menus. These results show clearly how important is to design the cross-device menus with a consistent order of items.

### 5.7. Potential future research

In our study, users transitioning between menus involved switching from a desktop web interface on a laptop device to a mobile application on a smartphone and then back again. Our study analyzed post-transitioning user performance regardless of the sequence in which the devices were used. We revisited the data and conducted a device-based analysis of user performance when transitioning to the laptop versus when transitioning to the mobile. We did not find any large statistical differences between user performance when re-locating items on the laptop versus on the mobile. However, further focused research is required to investigate possible significant differences in user performance when transitioning from and to different platforms. Our study also investigated only menus that have consistent labels across devices; however, potential future research could investigate cross-device menus with inconsistent labeling. Further, the effect of menu layout in our study was investigated with menus that applied only a consistent order of items, leaving a need for a further study on the effect of layout when menus have an inconsistent order of items.

Our user study examined user performance on cross-device menu interfaces in a specific interaction mode in which the user switches between devices frequently in a very short period, which is a common cross-device interaction mode. However, when an immediate transition does occur, it implies the target device has capabilities (e.g., screen size, input) that are missing or inferior on the source device. Thus, in some cases, these capabilities may affect menu-selection tasks. Therefore, future research should investigate the effects of device characteristics on post-transitioning user performance. In addition, further studies should explore post-transitioning user performance with multi-device menu interfaces on other cross-device interaction modes (e.g., simultaneous interaction mode), and when the time between transition from one device to another is long.

## 6. Conclusions

The primary contribution of this study is that it highlights the effects of a set of independent variables on user performance for different designs of cross-device menu interfaces. We conducted two experiments in this study. In Experiment 1, the variables of menu configuration, menu length, and menu item order significantly affected post-transitioning user performance. For example, users performed better with redundant menus, short menus, and menus designed with consistent item order than they did with complementary menus, medium menus, and menus with inconsistent item order. Another example of the important effects found is that menus with inconsistent item order had a significantly higher level of selection error compared with menus that had a consistent item order. Menus that had consistent item order had a more efficient learning curve and required lower NASA-TLX workload from users compared with menus that had inconsistent item order. We also identified a set of interaction effects between independent variables that greatly influenced post-transitioning user performance. For example, menu length and menu item order interacted significantly with each other, suggesting a lower variance between selection times of medium and short menus when operating consistent menus compared with inconsistent menus. Another result from Experiment 1 was that configuration × length has a smaller effect on post-transitioning user performance when the order of menu items is consistent. Thus, designers can construct cross-device menus with any combination of configuration (redundant or complementary) and length (short or medium) as long as menus have a consistent order of items.

In Experiment 2, we investigated the effect of menu layout (horizontal–vertical and vertical-vertical) with different menu lengths (short and medium) on post-transitioning user performance. We found no significant effect of layout, but a significant main effect of length and a two-way interaction effect of layout and length. However, the sizes of all effects were small because all the menus investigated in Experiment

2 applied a consistent order of items. Therefore, designers can choose any combination of menu layout (horizontal–vertical or vertical-vertical) and menu length (short or medium) across devices as long as the menus applies a consistent order of items.

Our study has theoretical implications because it provides the foundations for future research on user interaction with cross-device menu interfaces. The study results also have important practical applications that can be used to design cross-device menus that support users to re-locate menu items efficiently and effectively when transitioning between devices.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Research data

The data used to support the findings of this study are restricted to protect privacy. Data can be made available upon request from Khalid Majrashi, majrashik@ipa.edu.sa for researchers who meet the criteria for access to the data.

## Declarations of interest

None.

## Vitae

Khalid Majrashi is a computer scientist with an interest in human–computer interaction, mobile and ubiquitous computing, and software engineering. He is an Assistant Professor in the Department of Information Technology of Institute of Public Administration, Riyadh, Saudi Arabia.

## References


Als, B.S., Jensen, J.J., Skov, M.B., 2005. Comparison of think-aloud and constructive interaction in usability testing with children. In: Proceedings of the Conference on Interaction Design and Children. ACM, pp. 9–16.

Alsumait, A., Al-Osaimi, A., AlFedaghi, H., 2009. Arab children's reading preference for different online fonts. In: Jacko, J.A. (Ed.), Human-Computer Interaction. Interacting in Various Application Domains. Springer, Berlin, Heidelberg, pp. 3–11 HCI 2009 [Lecture Notes in Computer Science].

Bailly, G., Lecolinet, E., Nigay, L., 2008. Flower menus: a new type of marking menu with large menu breadth, within groups and efficient expert mode memorization. In: Proceedings of the Working Conference on Advanced Visual Interfaces. ACM, pp. 15–22.

Bailly, G., Oulasvirta, A., Brumby, D.P., Howes, A., 2014. Model of visual search and selection time in linear menus. In: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems. ACM, pp. 3865–3874.

Byrne, M.D., 2001. ACT-R/PM and menu selection: applying a cognitive architecture to HCI. Int. J. Hum. Comput. Stud. 55, 41–84.

Byrne, M.D., Anderson, J.R., Douglass, S., Matessa, M., 1999. Eye tracking the visual search of click-down menus. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 402–409.

Callahan, J., Hopkins, D., Weiser, M., Shneiderman, B., 1988. An empirical comparison of pie vs. linear menus. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 95–100.

Card, S.K., 1982. User perceptual mechanisms in the search of computer command menus. In: Proceedings of the Conference on Human Factors in Computing Systems. ACM, pp. 190–196.

Chae, M., Kim, J., 2004. Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones. Behav. Inf. Technol. 23 (3), 165–181.

Cockburn, A., Gutwin, C., Greenberg, S., 2007. A predictive model of menu performance. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 627–636.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates, Hillsdale, NJ.

Dees, W., 2011. Usability of nomadic user interfaces. In: Jacko, J.A. (Ed.), Human-Computer Interaction: Towards Mobile and Intelligent Interaction Environments. Springer, pp. 195–204.

Denis, C., Karsenty, L., 2004. Inter-usability of multi-device systems: a conceptual framework. In: Ahmed, S., Homa, J. (Eds.), Multiple User Interfaces: Cross-Platform Applications and Context-Aware Interfaces. Wiley, New York, pp. 373–384.

Dillon, A., Watson, C., 1996. User analysis in HCI—the historical lessons from individual differences research. Int. J. Hum. Comput. Study. 45 (, 6), 619–637.

Egan, D.E., 1988. Individual differences in human-computer interaction. In: Helander, M. (Ed.), Handbook of Human-Computer Interaction. Elsevier, pp. 543–568.

Forrester Research. (2013). The European tablet landscape. A technographics data essentials document. http://www.forrester.com/European?Technographics?Consumer?Technology?Online?Survey?Q4?2012/-/E-SUS1711-/The?European?Tablet?Landscape/quickscan/-/E-RES91561(accessed 30 May 2018).

Google. (2012). The new multi-screen world: uUnderstanding cross-platform consumer behavior. http://services.google.com/fh/files/misc/multiscreenworld_final.pdf(accessed 30 May 2018).

Google. (2018). Material design. https://material.io/guidelines/(accessed 12 May 2018).

Green, C.D. (1999). Classics in the history of psychology. http://psychclassics.yorku.ca/ (accessed 30 Novemebr 2017).

Halverson, T., Hornof, A.J., 2008. The effects of semantic grouping on visual search. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. ACM, pp. 3471–3476.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical research, Advances in Psychology. Elsevier, pp. 139–183.

Hornbæk, K., 2006. Current practice in measuring usability: challenges to usability studies and research. Int. J. Hum. Comput. Stud. 64 (, 2), 79–102.

Hornof, A.J., Kieras, D.E., 1997. Cognitive modeling reveals menu search in both random and systematic. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 107–114.

Jokela, T., Ojala, J., Olsson, T., 2015. A diary study on combining multiple information devices in everyday activities and tasks. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 3903–3912.

Kacmar, C.J., Carey, J.M., 1991. Assessing the usability of icons in user interfaces. Behav. Inf. Technol. 10 (, 6), 443–457.

Kaptelinin, V., 1993. Item recognition in menu selection: the effect of practice. In: Proceedings of INTERACT'93 and CHI'93 Conference Companion on Human Factors in Computing Systems. ACM, pp. 183–184.

Kara, P., Raluca, B. (2016). Hamburger menus and hidden navigation hurt UX metrics. https://www.nngroup.com/articles/hamburger-menus/(accessed 12 May 2018).

Lazar, J., Feng, J.H., Hochheiser, H., 2010. Research Methods in Human-Computer Interaction. John Wiley & Sons.

MacKenzie, I.S., 1992. Fitts' law as a research and design tool in human-computer interaction. Hum. Comput. Interact. 7 (, 1), 91–139.

Majrashi, K., 2016. Cross-platform user experience (Ph.D. Doctoral thesis). RMIT University.

Majrashi, K., Hamilton, M., Uitdenbogerd, A.L., 2016a. Correlating cross-platform usability problems with eye tracking patterns. In: Proceedings of the 30th British Human Computer Interaction Conference (HCI 2016). ACM, pp. 40–51.

Majrashi, K., Hamilton, M., Uitdenbogerd, A.L., 2016b. Cross-platform cross-cultural user experience. In: Proceedings of the 30th British Human Computer Interaction Conference (HCI 2016). ACM, pp. 20–32.

Majrashi, K., Hamilton, M., Uitdenbogerd, A.L., 2017. The ubiquitous device transition experience of the mobile user. In: Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. ACM, pp. 1–2.

McDonald, J.E., Stone, J.D., Liebelt, L.S., 1983. Searching for items in menus: the effects of organization and type of target. In: Proceedings of the Human Factors Society Annual Meeting. SAGE Publications, pp. 834–837.

Michelsen, C.D., Dominick, W.D., Urban, J.E., 1980. A methodology for the objective evaluation of the user/system interfaces of the MADAM system using software engineering principles. In: Proceedings of the 18th Annual Southeast Regional Conference. ACM, pp. 103–109.

Microsoft. (2013). Cross-screen engagement: multi-screen pathways reveal new opportunities for marketers to reach and engage consumers. https://advertising.microsoft.com/en/wwdocs/user/display/cl/researchreport/1932/global/Microsoft-Advertising-Insights-Cross-Screen-Engagement-White-Paper.pdf(accessed 31 May 2018).

Nielsen, J., 1994. Usability Engineering. Elsevier, San Francisco, CA, USA.

Nilsen, E.L., 1991. Perceptual-Motor Control in Human-Computer Interaction (Doctoral Dissertation). University of Michigan.

Pasqualotto, E., Simonetta, A., Gnisci, V., Federici, S., Belardinelli, M.O., 2011. Toward a usability evaluation of BCIs. Int. J. Bioelectromagn. 13, 121–122.

Pyla, P.S., Tungare, M., Pérez-Quinones, M., 2006. Multiple user interfaces: why consistency is not everything, and seamless task migration is key. In: Proceedings of the CHI Workshop on the Many Faces of Consistency in Cross-Platform Design. Montreal, Quebec, Canada.

Ramadan, M.Z., 2011. Evaluating college students' performance of Arabic typeface style, font size, page layout and foreground/background color combinations of e-book materials. J. King Saud Univ. Eng. Sci. 23 (, 2), 89–100.

Sauro, J. (2013). How to measure learnability. https://measuringu.com/measure-learnability/(accessed 15 Novemebr 2017).

Shin, D.-H., 2016. Cross-platform users' experiences toward designing interusable systems. Int. J. Hum. Comput. Stud. 32 (, 7), 503–514.
Somberg, B.L., 1987. A comparison of rule-based and positionally constant arrangements of computer menu items. ACM SIGCHI Bulletin 18 (, 4), 255–260.
Tracy, J.P., Albers, M.J., 2006. Measuring cognitive load to test the usability of web sites. In: Proceedings of the Annual Conference-society for technical communication, pp. 256.
Wäljas, M., Segerståhl, K., Väänänen-Vainio-Mattila, K., Oinas-Kukkonen, H., 2010. Cross-platform service user experience: a field study and an initial framework. In: Proceedings of the 12th International Conference on Human computer Interaction with Mobile Devices and Services. ACM, pp. 219–228.
Yang, H.-H., Chen, Z.-N., Hung, C.-W., 2017. Performance of smartphone users with half-pie and linear menus. Behav. Inf. Technol. , 36 , 9, 1–20.

A photograph of the author.