Breadth *and* depth: A comparison of search performance in hierarchical and mega menus

Samantha Naylor

Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Brain Sciences, University College London, 2016.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr Duncan Brumby, for his advice, guidance and encouragement through the project. Also, to my personal tutor, Dr Nic Marquardt for his calm, pragmatic and unswerving support. In addition, I would like to thank all the participants who gave their time willingly and cheerfully. A particular "thank you" goes to Mr Aaron Ashworth for his time, dedication and considerable coding skills in building the application. I would also like to extend a big thank you to my family and close friends who have offered love, patience and chocolate throughout my studies.

ABSTRACT

Menus are the ubiquitous method of navigating websites. Hierarchical (or "drop-down") menus are a very common menu design. A core design trade-off of hierarchical menus is between breadth and depth: a broad menu has fewer levels but more options to scan at each level, whereas a deep menu has more levels to traverse but fewer options at each level. Prior work has found that breadth is preferable to depth, except when there is a clear path to the target. Advancements in technology have now made possible mega menus that afford both depth and breadth. Mega menus are large panels that appear when hovering over a top-level navigation option. They display many lower level links, usually grouped into related categories, allowing the user to bypass intermediate levels and directly select a target item. An experiment was designed to investigate whether mega menus improve search performance. Sixteen participants completed two types of search (with and without a given route to the target) in both a hierarchical menu and a mega menu. Results showed that the type of menu did not affect search performance, which is inconsistent with previous studies that suggested mega menus would reduce search times. Participants were more likely to abandon searches when they were given only the target item, especially when using a hierarchical menu, which suggests that mega menus may reduce dropout rates on sites where visitors do not have a clear route to the target.

CONTENTS

CHAPTER 1 - INTRODUCTION

Interacting with websites is a task performed by millions of people every day. A Google.com search, for example, is typically followed by searching for content on one or more websites. As technology facilitates increasingly bigger and more complicated sites, it becomes even more important that users can find what they are looking for quickly and easily. This also has clear business benefits, such as increasing the likelihood that the visitor will engage with, buy from and return to a site (Nielsen, 2000). Conversely, the frustration caused to users who are unable to find what they are looking for can affect brand perception and make it more likely that they will go elsewhere (Krug, 2005).

Menus are the ubiquitous method of navigating websites. A menu is a series of linked pages and the structure of these linked pages, often called the “information architecture”, plays a key role in a site’s usability. As a result, over the past 20 years menu design, and the factors influencing it, have received much empirical and theoretical attention (Kiger, 1984; Kreigh, Pesot & Halcomb, 1990; Landauer, 1987; Larson & Czerwinski, 1998a; MacGregor, Lee & Lam, 1986; Miller & Remington, 2004; Norman, 1991; Parkinson, Sisson, & Snowberry, 1985; Schultz & Curran, 1986; Zaphiris, 2000).

Early menus tended to be hierarchical due to limitations in screen size and resolution and, despite advances in technology, hierarchical menus are still very common today. A hierarchical (or "drop-down") menu is a multi-level arrangement of options. Selecting a top-level item will reveal its sub-headings, and so on, until the lowest level of the hierarchy is reached (see Figure 1). Labels within the structure usually move from general descriptors at the top of the hierarchy, through increasingly specific category and sub-category descriptors until the lowest level is reached, from where the user can select their desired "target" item (Miller, 1981). Users are unable to view lower-level items without systematically going through the higher levels, selecting the most appropriate-looking option at each level.

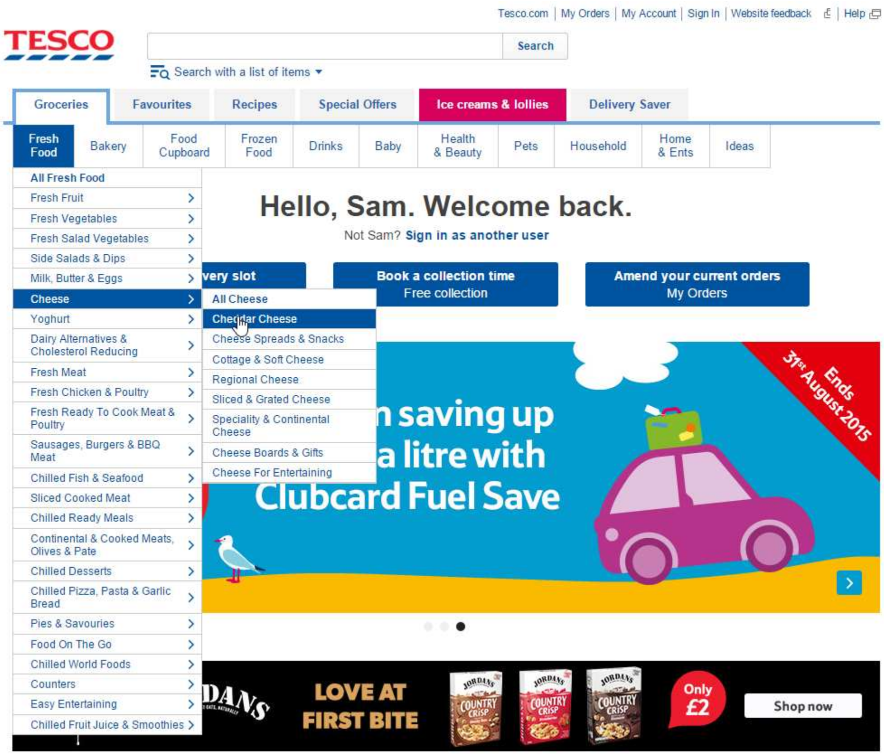


Figure 1: Example of hierarchical menu of an online retailer

One key aspect of these menus is the number of items per level (breadth) and the number of levels (depth). Consider for a moment a small website of 30 pages. There are many options for how to structure the site. At one extreme, each page could be given a link at the top level with no further levels (30^1), resulting in a very broad structure (Figure 2):

Figure 2: Example of broad structure of a site with 30 pages

At the other extreme, there could be just two options at each level (2^4), resulting in a much deeper structure (Figure 3). Intermediate options could present a more balanced hierarchy (e.g. 3^3).

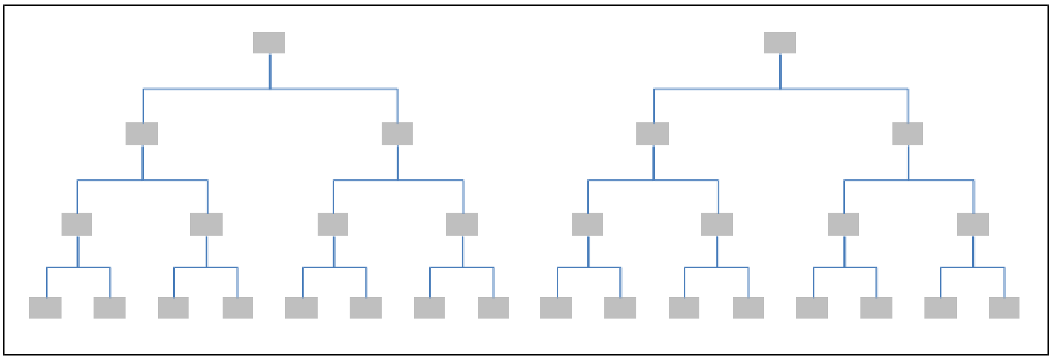


Figure 3: Example of a deep structure of a site with 30 pages

The optimal structure of these menus is referred to as the "breadth/depth trade-off". Broad structures have fewer levels to traverse but more options at each level. Fewer levels mean fewer decision points, and being able to see all options makes it more likely that a user will select the most appropriate route (Jacko & Salvendy, 1996). This minimises "path errors" and, if they do occur, reduces the time taken to recover (Landauer, 1987; Schultz & Curran, 1986). In addition, being able to visually scan all options has been shown to minimise short-term memory load as it reduces the need to remember the hidden sub-categories (Schultz & Curran, 1986). However, increased breadth means more options to scan at each level. This results in greater demands on visual search processes, which may increase the time taken to make a decision (Katz & Byrne, 2003). Conversely, deep menus have more levels to traverse, but fewer options to scan at each at level. In addition, deeper, more compact menus may require less cursor movement, thus decreasing selection time (Norman, 1991).

A modern design alternative that offers both breadth and depth is the "mega menu". Mega menus are large navigation panels that drop down or fly out from the top-level navigation revealing lower levels of the site's structure (Figure 4):

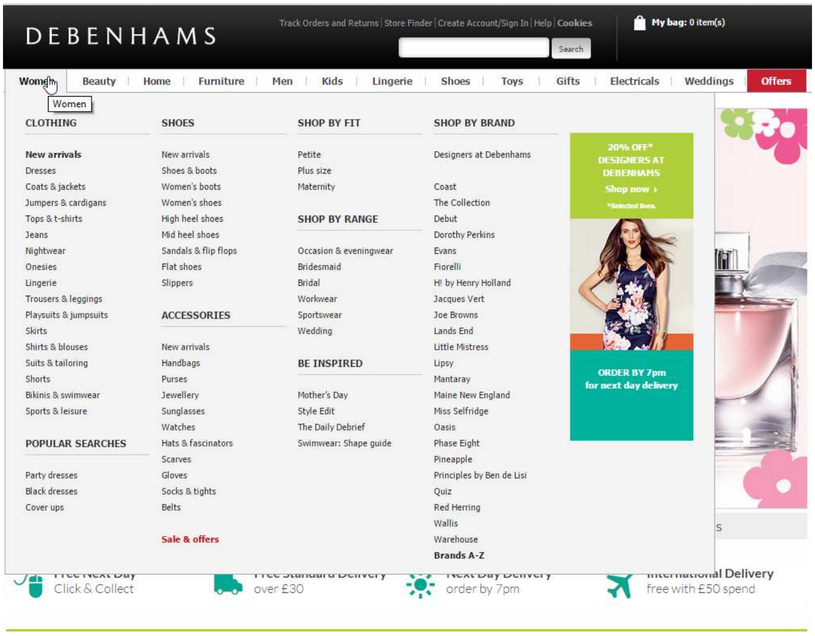


Figure 4: Example of a mega menu from an online retailer

Mega menus are able to display many options at each level (breadth) and share many similarities with broad hierarchical menus, but with two key differences. Firstly, the large navigation panel enables lower level links (depth) to also be displayed, so users can see the target items within each sub-category. Secondly, the user is able to bypass intermediate levels, if they so choose, and select lower level links without having to go systematically through each level of the hierarchy.

As mentioned, optimal menu design has been the subject of extensive research (e.g. Kiger, 1984; Kreigh et al., 1990; Landauer, 1987; Larson & Czerwinski, 1998a; Schultz & Curran, 1986; Tullis, 1985; Zaphiris, 2000). Whilst many of these studies were conducted prior to the emergence of mega menus, some have investigated aspects of menu design that foreshadow them. For instance, the “labelled” menus in Hornof and Halverson’s study (2003) are similar to mega menus. Their results showed that labelled layouts were searched faster than unlabelled layouts. However, Hornof and Halverson’s study was more concerned with search strategies and categorical grouping/labelling than the design of the menu itself. This present study aimed to explicitly investigate whether mega menus reduce search times. Results are important as they may lead to insights to assist designers in creating optimal website menus and, consequently, improve users’ search experience.

The next chapter will review existing literature, which highlights three main aspects to consider in relation to the performance of mega menus. Firstly, as previously mentioned, the breadth/depth debate is central. It will review evidence for the superiority of both breadth and depth and consider key papers that are particularly pertinent to the current study. Secondly, the type of task being undertaken will be considered as previous research (MacGregor et al., 1986; Miller & Remington, 2004; Zaphiris, Shneiderman & Norman, 2002) has demonstrated that having a clear path to the target item is a key factor in whether a broad or a deep structure is preferable. Thirdly, location learning will be reviewed as previous research (Larson & Czerwinski, 1998a) suggests that the optimum menu design may change as users become more familiar with the menu.

Chapter 3 describes a study in which participants searched for target items in a hierarchical menu and a mega menu. Two different types of search task were undertaken: searches where only the target item was given, and searches where the participant was given the target item and the route to get there. Data was captured for the first and second half of the experiment. The aim was to identify which menu design was most effective, the influence of the type of task being undertaken and how the optimum menu design changed with practice.

Chapter 4 presents the results of the experiment.

Chapter 5 provides a general discussion of the results of the experiment, their implications and their relation to previous research. It includes a reflection on the limitations of the study, a discussion of future research directions and, finally, gives a summary of the overall study.

CHAPTER 2 - BACKGROUND

This chapter considers three key aspects that have been shown to influence optimum menu design: the breadth/depth trade-off, search type and location learning. It concludes with a presentation of the research question and hypotheses.

Breadth/depth trade-off

Previous research has consistently shown that breadth is preferable to depth in a site's menu structure (e.g. Kiger, 1984; Landauer, 1987; Larson & Czerwinski, 1998a; Norman, 1991; Zaphiris, 2000; Zaphiris et al., 2002). For example, Miller (1981) conducted one of the earliest studies investigating how the structure of the menu hierarchy affects the speed and accuracy of target acquisition on a computer terminal. Participants completed search tasks on four different hierarchical structures, each having 64 items at the lowest level. The deepest condition consisted of binary choices at six different levels (2^6), the broadest displayed all 64 items in one large menu (64^1) and two middle options presented more balanced structures (4^3 and 8^2). Miller concluded that breadth was preferable to depth, as the two-level hierarchy produced the fastest search times and fewest errors.

A weakness of the study was that, whilst the deeper arrays arranged items into semantic categories, the broadest design (64^1) presented items randomly. It has been consistently shown that semantic categorisation of items facilitates search (e.g. McDonald, Stone & Liebelt, 1983, October; Parkinson et al., 1985; Vandierendonck, Van Hoe & De Soete, 1988), so this may have affected the results. Snowberry, Parkinson and Sisson (1983) repeated Miller's experiment to resolve this discrepancy and found that the broadest arrays (64^1 and 8^2) produced more accurate and faster search times than the deep arrays. As mega menus have fewer levels (like broad menus), it could be expected that they will reduce search times.

Various other studies have focused on different aspects of menu navigation, but also demonstrated the superiority of breadth in menu design. For instance, Kiger (1984) looked at different menu structures in a videotext-type system (2^6, 4^3, 8^2, 4x16 and 16x4) and found that the (broad) two-tier structure produced the fastest search times. More recently, Zaphiris's study (2000) confirmed that Kiger's findings were also applicable to web links. Zaphiris' subsequent study (2001) found that broad hierarchies were beneficial for systems catering for different age ranges as they resulted in optimum performances and smaller age-related differences.

Whilst mega menus did not exist when the majority of these breadth/depth studies were undertaken, two key papers consider aspects that foreshadow mega menus and so are worth reviewing. Parkinson, Hill, Sisson and Viera (1988) investigated the influence of breadth and depth in menu search performance using six menu configurations. The first three were constructed from the hierarchy used by Miller (1981) and Snowberry et al. (1983). In the next two menu designs, upon selecting an option the next two levels were displayed (rather than just one level) so that users could see the consequences of their choices. One option forced participants to choose a lower-level option, whilst the other forced them to make a selection at both levels. The final menu displayed lower levels and but users could choose how to respond: either selecting an option at each level, or bypassing levels. This final design has similarities with mega menus, which display lower levels and allow users to choose whether to bypass intermediate levels. It was this last design that yielded the best performance.

Leuthold, Schmutz, Bargas-Avila, Tuch and Opwis (2011) compared “vertical” and “dynamic” menus. In the former, higher-level navigation labels and target terms were always visible (similar to a mega menu). In the “dynamic” menus, users had to click on a category to display the target items it contained (similar to a hierarchical menu). “Vertical” menus resulted in better performance and higher subjective preference. Both of the above studies (Leuthold et al., 2011; Parkinson et al., 1988) suggest that mega menus will produce faster search times than hierarchical menus.

As can be seen, the majority of research has concluded that breadth is superior to depth in the design of menu structure. However, it is also worth considering three key studies that present an alternative view and point to two important factors that influence optimal menu design: search type and location learning.

Search type

Search type refers to the type of task being undertaken when searching a website's menu. A "simple" search task is one that puts little demand on the user, for example, searching for a single, distinct item with a clear route to the target (e.g. the "Titanic" DVD). Conversely, a "complex" search task is one that involves more cognitive effort; for example, searching for multiple items, or where the target item is not explicit (e.g. a DVD of a film starring Kate Winslet that is over five years old). "Simple" tasks have been shown to result in shorter search times than "complex" tasks. For instance, Brumby, Cox, Chung and Fernandes (2014) compared search performance between a known-item target (where participants were told exactly what to look for) and a semantic search (where only the target category was given). Known-item searches were faster and more accurate.

Studies have consistently shown that search type affects optimal menu design (e.g. Brumby et al., 2014; Leuthold et al., 2011). For instance, Hochheiser and Shneiderman (2000) compared simple and complex search tasks on simultaneous (similar to mega menus) and sequential (similar to hierarchical) menus. A "simple" task required one selection from each of three menus. A "complex" task required the participants to compare two selections and then use this information to make a selection from the third menu. Simple tasks were completed faster on sequential menus, whilst the simultaneous design was better for complex tasks.

In another study, Miller and Remington (2004) used a computational model to simulate how users navigate websites of two and three tiers. They concluded that a deeper (three-tier) structure was optimal but only when high-quality link labels made the path to the target clear. They reasoned that a clear route to the target reduced cognitive effort and, therefore, made the task "simpler".

The present study aimed to compare search performance between hierarchical menus and mega menus. As can be seen, optimal menu design may be affected by the type of search task being undertaken - "simple" or "complex" – so participants completed two types of search task. In a "target and route" (TR) search, participants were given the target item and the route to get there, i.e. a "simple" search with a clear route to the target. In "target only" (TO) searches, only the target item was given, i.e. a more "complex" search where participants were required to assess the options at each level and select the one they thought most likely to lead to the target.

As more complex search tasks will necessarily require greater cognitive effort, it would be expected that they would result in longer search times (Leuthold et al., 2011). Furthermore, based on these previous studies (Hochheiser & Shneiderman, 2000; Miller & Remington, 2004), it is expected that TR searches will be performed faster in the hierarchical condition and TO searches will be performed faster on mega menus. This is because these studies (Hochheiser & Shneiderman, 2000; Miller & Remington, 2004) found that a deeper structure was faster for simpler tasks (e.g. with a clear route to the target) and complex tasks were performed better on menus where users were able to scan the available options (e.g. a broad structure).

Location learning

With practice and a consistent layout, users remember where items in a menu are located (Card, 1982; McDonald et al., 1983). This is known as "location learning" and has been found to reduce search times in a visual search environment (Kreigh et al., 1990; Paap & Roske-Hofstrand, 1986; Schultz & Curran, 1986; Teitelbaum & Granda, 1983). For example, Paap and Roske-Hofstrand (1986) found that, with extended practice, users learnt the location of menu options and consequently, directly located the target resulting in shorter search times. The current study is interested in whether the optimal menu structure changes with practice, as participants "learn" where items are located.

In MacGregor, Lee and Lam's study (1986), participants completed search tasks on menu structures with 2, 4, 8 and 16 alternatives. Results showed that those new to the interface performed better on a deeper menu (with four or five alternatives per level). Hierarchical menus are similar to the "deep" menu in the above study (MacGregor et al., 1986) as they have more levels. This suggests that when participants are new to the menu, they will perform better on a hierarchical menu.

However, as users become familiar with the menu, the number of items they examine reduces, thus reducing decision time (Paap & Roske-Hofstrand, 1986).

MacGregor et al. (1986) continued by suggesting that as decision time decreases, optimal menu breadth increases. This implies that as users' search times reduce with practice, the (broader) mega menus will become the more efficient menu design.

In the current study, each condition was completed once by each participant, but half of the participants completed a given condition in Block 1 (the first half of the experiment), and half in Block 2. This enabled a comparison between those new to the interface, and those with more experience of the using the menu. Following MacGregor et al.'s (1986) findings, it would be expected that search times will decrease with practice and participants will perform searches faster on mega menus in Block 2 relative to hierarchical menus, i.e. whilst search times would be expected to decrease in Block 2 for both menus, the difference will be more pronounced in mega menus.

In summary, this study aimed to investigate whether mega menus offered search performance benefits over hierarchical menus. Prior research suggests three key aspects to be considered. Firstly, the compromise between breadth and depth is important, as previous studies have found that breadth is preferable to depth in a menu structure, except when there is a clear route to the target. With a clear route to the target, deep structures are faster. Secondly, search type impacts search performance with "simple" tasks performed quicker than "complex" tasks. Search type may also affect optimal menu design, with "simple" tasks favouring a deep structure and broader structures being more efficient for “complex” tasks. Finally, location learning suggests that, as users get familiar with a menu, they will find targets quicker and, after practice, a broad structure will be more efficient.

Summary and hypotheses

This study compared search performance between a hierarchical menu and a mega menu. Participants searched for target items with (“simple”) and without (“complex”) a given route to the target, in both menu designs. The hierarchical menu required participants to systematically select an option at each of three levels, whereas the mega menu displayed the two lower levels simultaneously and allowed participants to select a lower level item, bypassing the intermediate level (sub-category) if they chose. To enable the effects of location learning to be investigated, 40 searches were completed in the first half of the experiment (Block 1) and 40 in the second (Block 2). At the end of experiment, participants were asked which menu type they preferred.

Six hypotheses are proposed, based on prior work:

1. Mega menus will produce faster search times than hierarchical menus. This prediction is based on studies by Miller (1981) and Snowberry et al. (1983), who found that broad structures (similar to mega menus) produced faster search times than deep structures (similar to hierarchical menus). In addition, Parkinson et al (1988) and Leuthold et al. (2011) found participants located

items faster in menus that closely resemble mega menus than hierarchical menu designs.

2. TR (“simple”) searches will be faster than TO (“complex”) searches. This is based on Brumby et al.’s (2014) study that found “simple” searches (i.e. for known items) were faster than “complex” searches (i.e. “semantic” searches).

3. With practice, search times will decrease. This is following previous research (Kreigh et al., 1990; Paap & Roske-Hofstrand, 1986; Schultz & Curran, 1986; Teitelbaum & Granda, 1983) which found that as users became familiar with an interface, they learnt the location of items, and this reduced search times.

4. Participants will perform TO searches quicker in mega menus than hierarchical menus. This prediction is also based on Hochheiser and Shneiderman’s (2000) study whose simultaneous design (similar to a mega menu) resulted in quicker search times for more complex (e.g. TO) tasks.

5. With practice, the advantage of hierarchical menus for TR searches will reduce. This is based on previous research (Hochheiser & Shneiderman, 2000; Miller & Remington, 2004) which found that whilst easier tasks (e.g. those with a clear route to the target) were completed faster on deep menu

structures (similar to hierarchical menus), this was only apparent when users were new to the interface (MacGregor et al., 1986).

6. The benefits of mega menus will increase with practice. This prediction follows MacGregor et al.'s study (1986) that over time, users learnt the location of items, which reduced search times and increased the optimal breadth of the menu.

The next chapter will describe an experiment designed to compare mega menus and hierarchical menus in a visual search task, and try to disprove these hypotheses.

CHAPTER 3 - EXPERIMENT

Method

Participants

Sixteen participants (10 female) were recruited from the author's professional contacts and via an existing database of usability testing volunteers. The age of participants ranged from 28 to 53 years (*M* = 41.6, *SD* = 7.51). Participants gave their time in exchange for the chance to win Amazon vouchers based on their search performance over all the conditions. All participants were experienced computer users, had normal or corrected-to-normal vision, were native English speakers and had full dexterity in their upper limbs.

Design

A 2 x 2 x 2 (*menu type x search type x block*) mixed-factorial within-subjects design was used. The main independent variable was the type of menu, with two levels: hierarchical menu and mega menu. In the hierarchical menu condition, participants selected a top-level option, then a sub-category and then a target item. In the mega menu condition, selecting a top-level category revealed a large panel containing all target items within that category. Target items were grouped by sub-category and the participant could directly select a target item.

The second independent variable was the type of search, with two levels: “target only” (TO) where only a specific, named target item at the lowest level of the menu was given (e.g. “cheese”); and “target and route” (TR) where the participant was given the target item and the route to navigate to it (e.g. “food > dairy > cheese”).

The third independent variable was block. The four conditions (*search type vs menu type*) were counter-balanced to appear in either the first half of the experiment (Block 1) or the second half (Block 2). This enabled learning effects to be investigated.

The primary dependent variable was user performance operationalised by the time taken to locate the target item. The second dependent variable was the number of trials that were abandoned. The third dependent variable was subjective preference obtained from qualitative data from semi-structured interviews (Appendix A).

Materials

Menu type

Participants were presented with either a hierarchical menu or a mega menu in which to locate the target item. In the hierarchical menu, only the first level (category) was initially displayed. Upon selecting a category, the second-level (sub-category) options were revealed. Upon selecting a sub-category, the 10-item list appeared, from which participants could select their target (Figure 5).

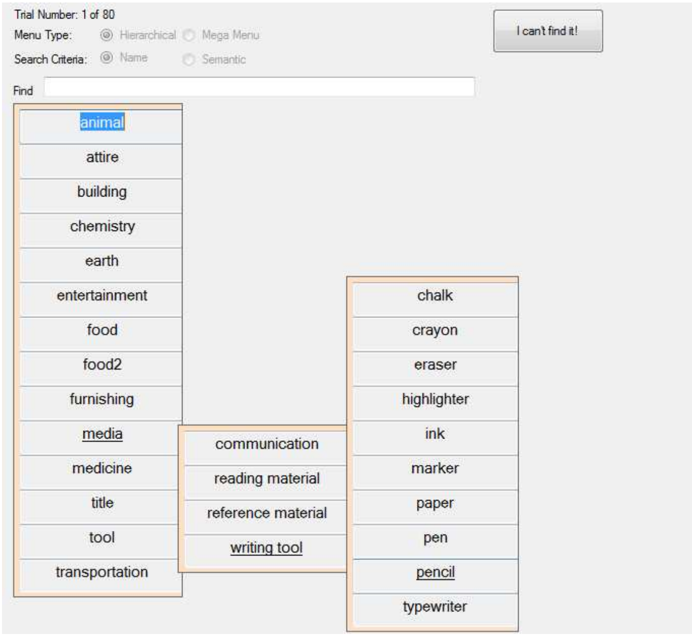


Figure 5: Example of hierarchical menu

In the mega menu condition, the top-level navigation looked identical to the hierarchical menu. However, upon hovering over an option, a large navigation panel appeared with four columns displaying both the sub-categories and their associated item lists. The sub-categories were highlighted and vertical dividers separated the semantic groups beneath them (Figure 6):

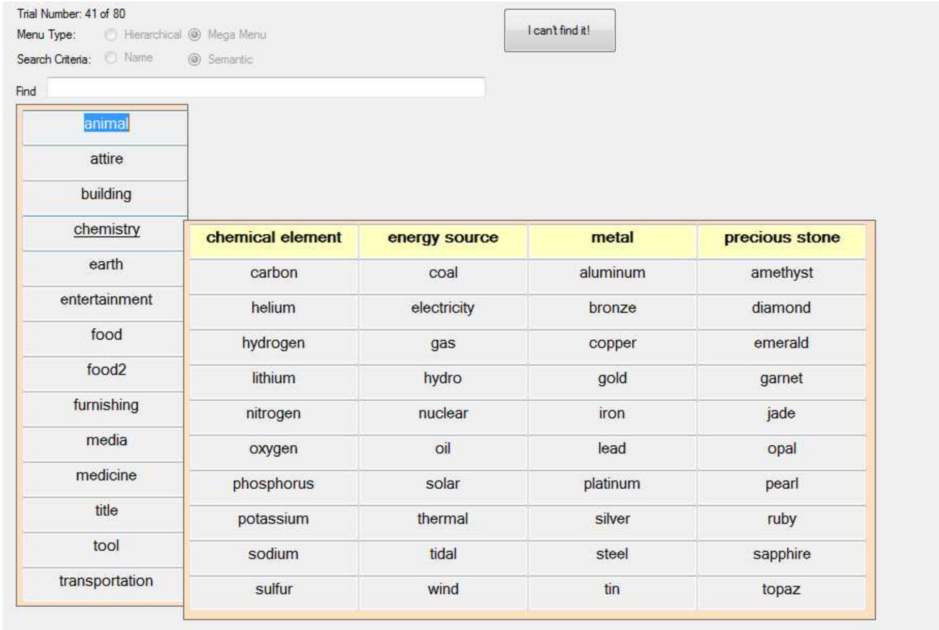


Figure 6: Example of mega menu

To accurately reflect common menu interaction, the menus remained visible when moving the mouse diagonally between sub-category and item targets, even if the mouse was momentarily not over a menu (to prevent a "hover tunnel"). Hovering over a target caused the target to become underlined, to reflect common website menu functionality. In addition, the mega menu and the hierarchical drop-downs rendered sensibly based on the window size, ensuring that all options were visible without scrolling. As a result, item lists from categories and sub-categories that were low in the design did not drop below the fold, but the whole list shifted upwards so all items were visible (see Figure 5 above).

Menu content

Menus were based on the same dataset of words used in previous studies (e.g. Bailly, Oulasvirta, Brumby & Howes, 2014; Brumby et al., 2014; Yoon et al., 2004). All menus consisted of 14 top-level categories (e.g. chemistry, entertainment, transportation) each of which had four sub-categories (e.g. chemical element, energy source, metal, precious stone are all members of the "chemistry" category). Within each sub-category, there was a list of 10 items (e.g., carbon, helium, hydrogen, lithium and nitrogen are all members of the "chemical element" category). In total, there were 560 unique words (with character length $M = 6.2$, $SD = 2.0$, range: 2 – 12).

The dataset had been previously validated with human participants to ensure items were grouped into semantically related categories (see Bailly et al. (2014) for details). Using the same dataset in all conditions removed any potential effects of the relevance of distractor items as this has been shown to be a confounding variable (e.g. Brumby & Howes, 2008). All target items were present in the menu (no trials involved searching for an absent target) and menu content was unfamiliar to all participants.

Because it is well known that target position impacts search performance (e.g. Byrne, Anderson, Douglass, & Matessa, 1999; Hornof & Kieras, 1999;

Vandierendonck et al., 1988), the position of the target item was controlled at each of the three navigation levels. Within each block of trials:

- half the trials accessed the target item via the first 7 categories (and half the last 7);
- half the trials accessed the target item via the first 2 sub-categories (and half the last 2);
- target items were located in each of the ten item positions twice in every group of 20 trials.

Menu design

Visual design has been found to impact search performance (Bailly et al., 2014; Brumby & Zhuang, 2015; Shneiderman & Plaisant, 2005), so the “look and feel” of the hierarchical and mega menus was kept as similar as possible (see Figure 5 and Figure 6). In the hierarchical menu, items in each list (category, sub-category and target items) were arranged in a single column (with a width of 150 pixels). The horizontal separation between items in the column was 25 pixels, with each item separated by a horizontal bar.

In the mega menu, the top-level navigation was identical to the hierarchical menu. Upon selecting a category, a large navigation panel was displayed. The navigation panel consisted of four columns (each with a width of 150 pixels). At the top of each column was a sub-category heading and its associated ten target items were arranged vertically underneath it. Sub-categories were in bold font, with a pale yellow background.

All text in both menu types was presented in an Arial font, size 10 and centred within the column. Categories, sub-categories and target item lists were all ordered alphabetically in every condition (so their positions remained constant throughout the experiment). This reflects common commercial menus, and allowed participants to learn the location of items.

Information video, consent form and post-experiment semi-structured interviews

Prior to the experiment, participants were shown a video (Appendix B) to explain the purpose of the experiment and what they would be asked to do. They were then given an instruction sheet and a consent form to sign (Appendix C). After completion of the 80 trials, the author conducted semi-structured interviews (Appendix A) to elicit preference information. Responses were captured manually in an Excel spreadsheet.

Equipment

A Dell Latitude E6410 laptop was used to run the test, which was built in C #, and run in “Visual Basic 2013 Express for Desktop”. The software enabled the experimenter to input participant ID and set the first menu type and search type to be presented. The software captured the test data throughout the experiment (see Appendix D).

Procedure

A pilot study was run to ensure a robust design and useful output data. Minor changes were made to the design of the mega menu (to make sub-categories clearer) and the instruction sheet was re-formatted into an instructional video to make it more accessible and easier to understand. In order to minimise the effects of fatigue, the experiment was designed to last only about 20 minutes. However, each session was booked for 30 minutes to ensure external time pressures did not affect results.

Sessions took place in a quiet, comfortable room where the participant would not be disturbed for the duration of the experiment. Participants were firstly thanked for their time and they then watched the instructional video. It explained that the aim of the experiment was to compare two different types of menu: hierarchical menus and mega menus. They were going to complete 80 searches (40 on each type of menu), which would take about 20 mins. They should be as quick and accurate as possible, and there were £50, £30 and £20 Amazon vouchers “up for grabs” for the winners. It then described the procedure they would follow.

After watching the instructional video, participants were given the opportunity to ask questions. Once they fully understood what was to take place, they were given the consent form to read and sign. Prior to the session commencing, the experimenter inputted their Participant ID into the application, and determined which menu type and search type would be performed first.

The order in which menus were presented was manipulated as a between-subjects factor: half of participants used the hierarchical menu first, and the other half the mega-menu. Half of the participants started with TO searches, and the other half with TR searches. This was counter-balanced with the presentation of menu type to mitigate any ordering effects (see Figure 7).

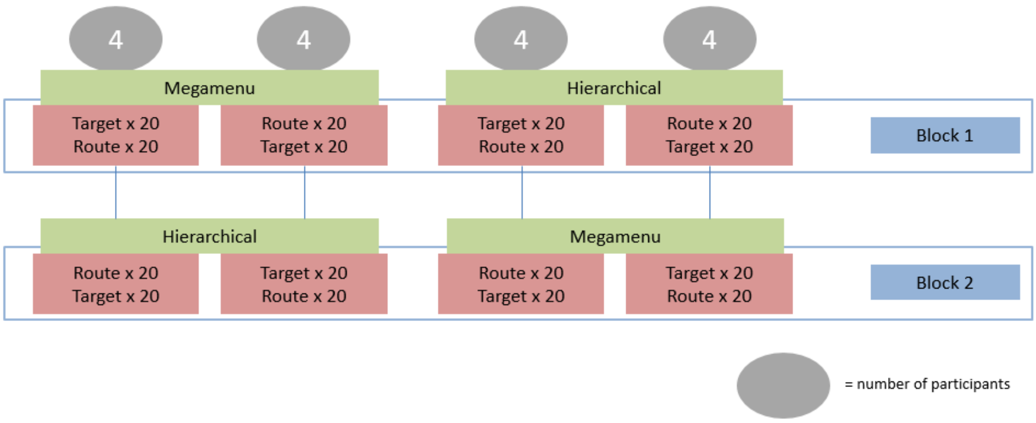


Figure 7: Four initial conditions and the order of subsequent conditions showing three independent variables: menu type (green), search type (red) and block (blue)

Participants starting with one search type on their first menu (e.g. TO search in a hierarchical menu) started with the other search type on their second menu (e.g. TR search in a mega menu), and vice versa. Participants were systematically allocated to start with one of the four conditions (Table 1).

	Hierarchical menu	Mega-menu
TO > TR	4 participants	4 participants
TR > TO	4 participants	4 participants

Table 1: Initial presentation based on menu- and search-type to mitigate ordering effects

When the participant indicated they were ready to begin, the experimenter clicked “Start Trial” and left the room. Forty trials were then completed on the first menu type (20 of each search type). The system presented the first search description to the participant (Figure 8).

Figure 8: Example of displayed search description of (a) “target only” search, and (b) “target and route” search

To reflect real-life navigation tasks, participants were required to remember the search description. Once they had memorised what they were to find, the participant clicked the “Go” button. They then used the displayed menu to locate the target item. Search time was defined as the time between the participant clicking the “Go” button and selecting the correct target item.

If the participant was unable to find a target item, they had the option of clicking the “I can’t find it!” button to abandon the trial. The system noted that the search had been abandoned. To simulate real-life search tasks, participants were dissuaded from clicking the "I can't find it!" button by adding two minutes to their total time for each abandoned search.

After completing 40 searches on the first menu type (20 of each search type), the system invited the participant to take a break in order to minimise fatigue. The length of their break was noted. Participants were incentivised based on total time to complete the 80 trials. This included search and pause times but excluded the break between menu types, so they could take as much time as they needed to ensure they felt rested.

When they were ready to start the second half, they clicked "Start 2nd half" which presented the first search description for the second menu type. The participant then proceeded to complete 40 trials (20 of each search type) on the second menu type. Once they had completed 40 searches on the second type of menu, they were presented with a thank-you message. It took approximately 20 minutes to complete all 80 trials.

The experimenter then returned to the room and conducted a semi-structured interview to gather qualitative data on participant preferences, frustrations and other aspects of the experience. Most had clear preferences and elucidated (often without prompting) on the reasons for these. It quickly became evident that menu preference in the TO search condition was dependent on how confident the participant was that that they could find the target item. Therefore, menu preference was captured for the following conditions:

- TR search
- TO search where the route to the target was unambiguous, i.e. the participant had a clear idea of the route to the target item
- TO search where the route to the target was ambiguous, i.e. the participant did not have a clear idea of the route to the target item

The experimenter recorded preferences and comments manually prior to the next participant. Participants were then thanked for their time and advised when the results of the competition would be announced.

Upon completion of all 16 participants, performance times were calculated and an email circulated (anonymously) to all participants. The top three performers received £50, £30 and £20 Amazon vouchers via email and were congratulated on their success.

CHAPTER 4 - RESULTS

For statistical analysis, a 2 x 2 x 2 (*menu type x search type x block*) repeated-measures ANOVA was used, with a .05 significance level for judging the significance of effects. Abandoned trials and trials in which the participant did not select the correct target item on the first attempt were excluded to enable a more accurate comparison between conditions.

Figure 9 shows selection time across the various experimental conditions. First, we consider the effect of search type. As can be seen in the figure, participants were faster in TR searches (M = 5.4 s, SD = 1.01 s) than TO searches (M = 11.23 s, SD = 2.73 s). Results of the ANOVA analysis found a significant main effect of search type on selection time, $F(1,15) = 89.47$, $p < .001$.

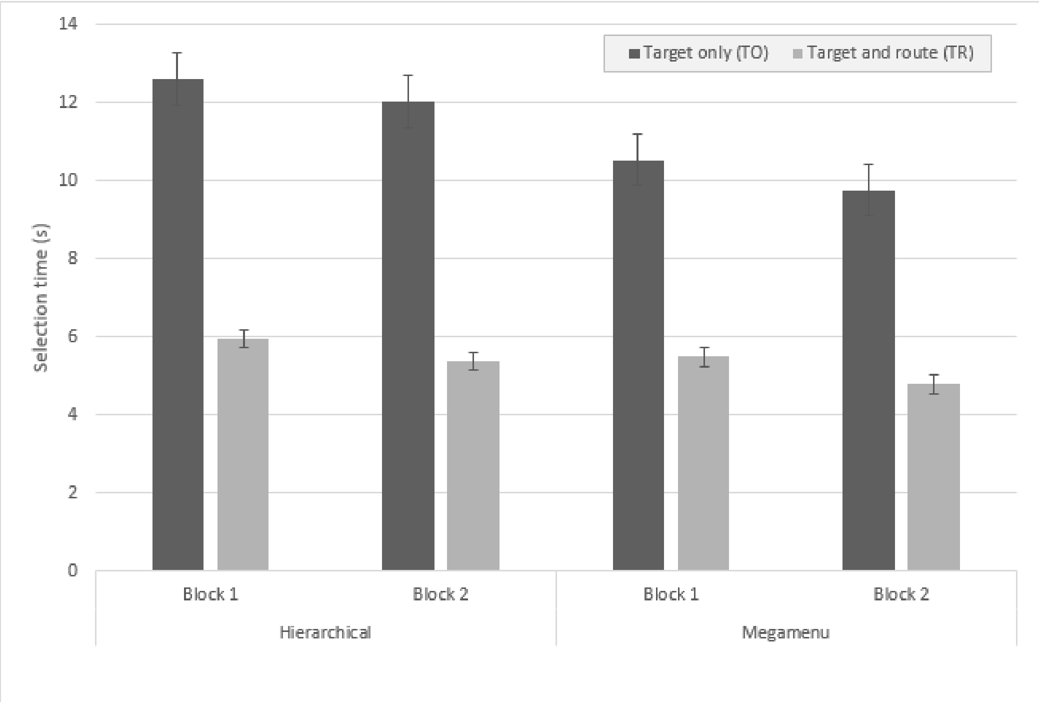


Figure 9: Selection time across the various experimental conditions

We next focus on the effect of menu type. Participants selected items faster in the mega menu (*M* = 7.64 s, *SD* = 2.36 s) than the hierarchical menu (*M* = 8.99 s, *SD* = 3.32 s), but the ANOVA analysis showed no significant effect of menu type on selection time $F(1,15) = 4.319$, $p < .055$.

Next, we consider the effect of block. As can be seen in Figure 9, participants were faster at selecting the target in Block 2 (*M* = 7.99 s, *SD* = 2.99 s) than Block 1 (*M* = 8.64 s, *SD* = 2.7 s) in all conditions. However, results of the ANOVA showed no significant main effect of block, $F(1,15) = 3.89$, $p = 0.67$. No interactions were significant.

Participants showed a preference for the mega menu, especially when given an ambiguous term in a TO search. A Chi-square test (Table 2) was performed to examine the relationship between search type and participants' reported menu type preferences. The relationship between these variables was significant (X^2 (4, N = 16) = 10.97, p = .027).

	TR	Unambiguous TO	Ambiguous TO
Hierarchical	6 (37.5%)	8 (50%)	2 (12.5%)
Mega menu	8 (50%)	5 (31.25%)	14 (87.5%)
None	2 (12.5%)	3 (18.75%)	0 (0%)

Table 2: Number of participants who preferred each menu type when performing the three types of search (TR, unambiguous TO and ambiguous TO). Numbers in brackets indicate column percentages.

Finally, we consider abandoned trials. Across all experimental conditions, 5% of trials were abandoned. Here we consider whether participants were more likely to abandon trials during different conditions. Figure 10 shows the percentage of trials abandoned across the various experimental conditions:

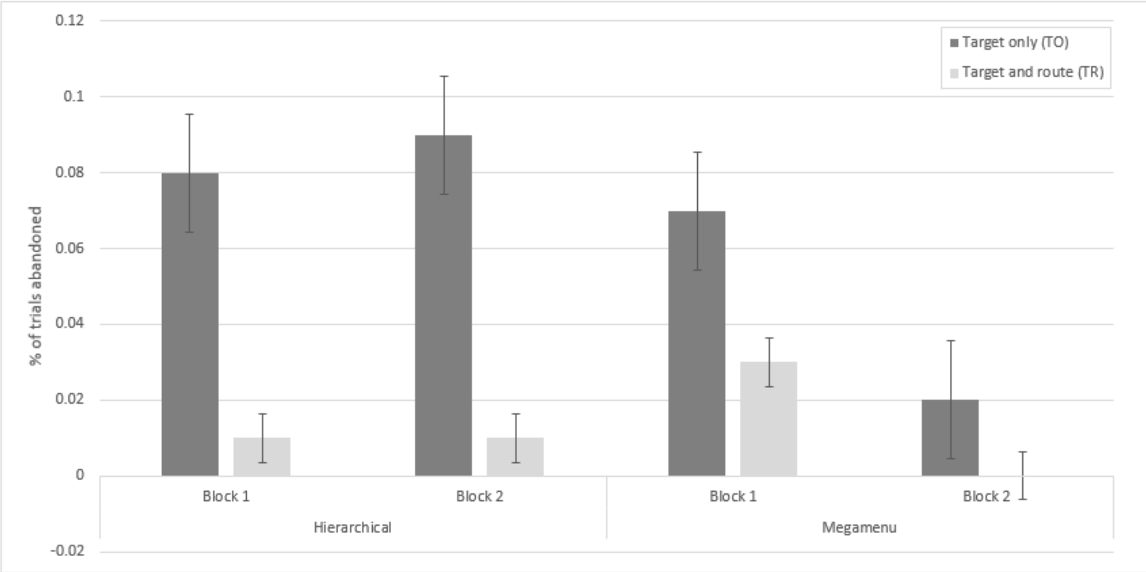


Figure 10: Percentage of trials abandoned across various experimental conditions

Considering the effect of search type, as can be clearly seen in the above figure, participants were much more likely to abandon TO searches (*M* = 0.07, *SD* = 0.1) than TR searches (*M* = 0.01, *SD* = 0.04). Results of the ANOVA showed a significant effect of search type, $F(1,15) = 14.22$, $p < .002$. There were no significant main effects of menu type or block.

Turning our attention to interaction effects, participants were more likely to abandon TO searches using a hierarchical menu (*M* = 0.09, *SD* = 0.12) than a mega menu (*M* = 0.05, *SD* = 0.08). The ANOVA showed a significant interaction effect between menu type and search type, $F(1,15) = 4.574$ $p < 0.05$. This supports qualitative feedback from the semi-structured interviews in which participants consistently mentioned that target-only searches were particularly problematic in the hierarchical menu condition.

In addition, participants were significantly more likely to abandon trials during Block 2 when using the hierarchical menu. The ANOVA showed a significant interaction effect of block and menu type, $F(1,15) = 6.607$, $p = 0.02$. No other interactions were significant.

CHAPTER 5 - DISCUSSION

Results of this study are now discussed and design implications considered, along with limitations of the study and suggestions for further research.

H1: Mega menus will produce faster search times than hierarchical menus

The results of this study show that there was no significant difference in search times between the mega menu and the hierarchical menu. This result is unexpected given previous work in this area. Prior research has consistently demonstrated the performance benefits of broad menus (similar to mega menus) over deep menus (e.g. Kiger, 1984; Landauer, 1987; Larson & Czerwinski, 1998a; Norman, 1991; Zaphiris, 2000; Zaphiris et al., 2002) and menus that are similar in design to mega menus (e.g. Leuthold et al., 2011; Parkinson et al., 1988). However, whilst participants selected items faster in the mega menu than the hierarchical menu in this experiment, the effect was not significant.

There are a number of potential reasons why the results did not support prior research. Firstly, the mega menu may not have been comparable to the broad menus of previous studies. In this experiment, the mega menu displayed 44 items, many more than the 16 items of broad menus in previous research (e.g. Kiger, 1984; Zaphiris, 2000). Increasing the display area and the number of items would have necessarily increased the number of items to be scanned and may have resulted in longer response times in the (broader) mega menu. Indeed, some participants reported greater cognitive effort was required when scanning the mega menu due to the large number of items to be reviewed. This may have increased search time and outweighed the benefit of fewer levels (than the hierarchical menu).

Secondly, the design of the mega menu may have affected results. Some participants reported that they did not notice the sub-categories on the mega menu immediately, so were searching the entire category dataset. Previous research has shown that visual grouping aids search performance (Bailly et al., 2014; Brumby & Zhuang, 2015), so this may have resulted in increased search times in the mega menu conditions for these participants. To improve the current study, the delineation of sub-category could be made more noticeable in the mega menu.

Thirdly, it is possible that the mega menu was not “mega” enough. Each top-level option had just four sub-categories, each with 10 items; however, a commercial website could easily have multiple times this number of lower-level links. For instance, returning to the example in Figure 1, the “Fresh food” top-level option (highlighted) contains 24 sub-categories; the sub-category “cheese” contains 8 sub-sub-categories; the sub-sub-category “cheddar cheese” 7 lower categories and “Mature cheddar cheese” contains 24 items. Translating this example to a hierarchical menu would result in more levels, potentially making the differences between the two menu types more pronounced. It would be interesting to repeat the experiment with a larger dataset to determine if and how it affects optimal menu design. Further to the above example, this would have real-world ecological validity.

H2: TR searches will be faster than TO searches

Results showed that participants performed TR searches quicker than TO searches. This finding supports previous research (Brumby et al., 2014) that demonstrated participants are faster at finding the target when the route to the target is clear (a "simple" search). In TR searches, participants were scanning the list of options for a known term. However, in TO searches participants would have had to make judgements on the relevance of the link label as to whether it would lead to their target item, which would necessarily take more cognitive effort, and more time (Miller & Remington, 2004). Therefore, it was expected that TR searches would be performed quicker in all conditions.

This could also explain the finding that participants were significantly more likely to abandon TO searches than TR searches. It is perhaps not surprising that participants were more likely to abandon searches that required greater cognitive effort.

H3: With practice, search times will decrease

Results of the study showed that participants did not get faster at locating targets with practice: whilst mean search time was faster in the second block of trials than the first (in all conditions), the difference was not significant. This result is inconsistent with previous research that suggested location learning reduced search time (e.g. Paap & Roske-Hofstrand, 1986). This may have been a result of the number of trials undertaken. Paap and Roske-Hofstrand (1986) found that users learnt the location of items “with extended practice”. This present experiment may have been too short for any learning effects to become apparent as it lasted only 20 minutes. Extending the experiment to include more trials may make any learning effect more noticeable. Interestingly, participants reported trials where they had seen the target word on a previous search and so were able find it very quickly, demonstrating that location learning was indeed taking place.

H4: Participants will perform TO searches quicker in mega menus than hierarchical menus

Results showed that participants did not perform TO searches faster in the mega menu than the hierarchical menu. This result was unexpected as previous research (Hochheiser & Shneiderman, 2000) had shown that broader menus (similar to mega menus) were superior for “complex” searches, where the route to the target was less clear. The result of this study could be explained by two unpredicted confounding variables.

Firstly, there was variation in the ambiguity of the target item. For instance, whilst participants could be fairly confident that “lion” would be under “animal” > “wild animal”, the target item “rock” could reasonably have been found under “earth”, “food” or “entertainment” (it was, in fact, in this last category, under the sub-heading "music genre"). For TO searches, participants reported a clear distinction between targets for which they had a high level of confidence of the route to the target and those for which they were less confident of the route.

Secondly, there was variation in the quality of the link labels, which would have affected participants' ability to find a target-only item. For instance, two of the top-level navigation options were "food" and "food2". From these descriptors, it is difficult to know which would lead to a food-related target item. "Food" and "food2" are low-quality link labels. In contrast, the top-level navigation option "building" is a high-quality link label, as there is little ambiguity in its meaning and so it gives a clearer indication of whether a target item may be found below it.

High-quality link labels create a clear path to the target, or a strong "information scent". A key concept in Information Foraging Theory (IFT), "information scent" refers to environmental (or "proximal") cues that are used to evaluate options when navigating. Users will choose the option that gives them the clearest indication (or strongest scent) that it will lead to their goal (Pirolli, 2003). Returning to Miller and Remington's (2004) study, only targets with unambiguous link labels were found faster in the three-tier (deeper) structure.

Both the ambiguity of the target item and the quality of the link labels may have affected results. Participants reported that TO searches for unambiguous items, or routes with high-quality link labels, were similar in complexity to TR searches.

Therefore, TO searches overall may not have been “complex” enough to see performances differences between the two menu designs.

If this study is repeated, it may be advantageous to make “simple” and “complex” searches more distinct. For instance, participants could either be told exactly what to look for (“simple”) or just given the dictionary definition of the target (“complex”), similar to Brumby et al.’s (2014) “known-item” and “semantic” searches. Alternatively, participants could either be given goal with a single criterion (e.g. “cheese”) or goal with many criteria (e.g. “something you might put in a salad, in a sandwich, or on toast”), similar to Leuthold et al.’s (2011) “simple” and “complex” tasks. This may make the relative performance of hierarchical menus and mega menus more pronounced.

H5: With practice, the advantage of hierarchical menus for TR searches will reduce

Previous research had shown that deep structures produced quicker search times when there was a clear route to the target (e.g. Hochheiser & Shneiderman, 2000; Miller & Remington, 2004) but only when the interface was unfamiliar to users (MacGregor et al., 1986). However, in this study participants did not perform TR searches significantly faster in the hierarchical menu when they were new to the interface (Block 1). In fact, TR search times did not differ substantially between menu designs throughout the experiment. This suggests that the two menu designs were more similar than expected.

One potential reason for this result is that the data structure between the two menus was the same. For instance, the route to “stomach” was “medicine > human organ > stomach” in both structures. In contrast, the data structure in previous studies was changed to create different breadth/depth conditions. For instance, in the 4^3 condition of Miller’s (1981) study, the route to “stomach” was biology > anatomy > stomach. However, the route in the 2^6 study traversed six levels. Insufficient detail is included in Miller’s (1981) paper, but it can reasonably be assumed that the structure of the data would have been different (to create six levels). In Miller’s (1981) study, greater differences in the number of levels may have made the differences between the menu structures more pronounced. On reflection, it is perhaps unsurprising that the differences in TR search times in the current study were less noticeable as the data structure was the same in both menu designs.

H6: The benefits of mega menus will increase with practice

There was no discernible improvement in participants’ search performance on the mega menu, relative to the hierarchical menu, in the second block of trials. Search times decreased in all conditions, but not significantly. Previous research (MacGregor et al., 1986) had suggested that, as location learning reduced decision-making time, the optimum breadth of the menu would increase. Therefore, it was expected that mega menus would be more efficient than hierarchical menus after practice (i.e. in Block 2). However, this was not observed in this study.

The current study involved a total of 80 trials and took most participants less than 20 minutes to complete. It is possible that this was too short for any learning effects to become apparent. As mentioned previously, it would be interesting to see whether there was a greater effect if participants were given time to get even more familiar with the structure, although care should be taken to ensure fatigue did not adversely impact results.

However, analysis of abandonment rates showed two interesting changes between the first and second block of trials. Firstly, while participants abandoned TR searches in a hierarchical menu at a consistent (and minimal) rate (0.01% in both blocks), abandonment rates for TR searches in a mega menu dropped from 0.03% to 0%: no participants abandoned a TR search on the mega menu in the second block of trials.

Secondly, for TO searches in a hierarchical menu, abandonment rates actually increased in the second half of the trial (from 0.08% to 0.09%). Conversely, TO searches in a mega menu dropped from 0.07% (Block 1) to 0.02% (Block 2) which suggests that participants got better at using the mega menu over time. Whilst there was no interaction effect of menu type and block when considering search performance, the interaction of menu type and block on abandonment rates was significant: with (minimal) practice, visitors were less likely to abandon their search on a mega menu, as opposed to a hierarchical menu. This finding was consistent irrespective of whether there was a given route to the target. Reduced abandonment rates could reasonably mean fewer visitors leaving a website in the real world so using a mega menu could translate into real business benefits.

Participant preference

Overall, participants preferred using the mega menu. This could be explained by considering the cognitive effort required to use the two menu designs. A deeper menu has more levels and, therefore, requires the user to make more decisions. As each level involves visual search, decision and response selection, the overall cognitive effort is greater for a deep menu than a broad menu where relatively fewer choices will have to be made (Jacko & Salvendy, 1996). In addition, deep structures may increase short-term memory load as users try to remember hidden sub-categories (Schultz & Curran, 1986). It is perhaps not surprising that users preferred the menu design that required less effort to use, even if there was no significant difference in search times.

Considering participant preference by search type, there was no distinct preference for TR searches. Similarly, for TO searches where participants were confident where to locate the item, menu preference was less pronounced (indeed, three participants reported no preference of menu type). However, when participants were less confident about where to find an item, there was a strong preference for mega menus (14 vs 2 participants), suggesting that they found it easier to explore a mega menu when they were unsure of where to find the target item. This has important implications for real-world sites, as visitors are less likely to know where to find what they are looking for, especially when accessing a website for the first time.

Limitations and suggested improvements

Five main limitations should be borne in mind when considering the results of this study. The first limitation has already been discussed: the unforeseen split of TO searches between ambiguous and unambiguous search terms. In addition, certain terms were culturally sensitive. For example, “pumps” (found via “attire” > “footwear”) is not a term commonly used in the UK; and a “trailer” (found via “building” > “home”) is not recognised as a building type, but a contraption that is towed by a motor vehicle. As all participants were British nationals, culturally-sensitive items may have distorted the results, especially for TO searches. To improve upon the current study, search times could be analysed and a card-sorting exercise conducted. Culturally-sensitive items could be removed and the remaining target items categorised into “unambiguous” and “ambiguous” searches. This factor could then be controlled, and the differences between “target and route”, “unambiguous target-only” and “ambiguous target-only” searches (and the interaction effects with menu type and block) investigated.

Second, whilst based on a previous study (Yoon et al., 2004), the dataset was not reflective of common online content through which a user would need to search. Real-world websites are likely to be focused on one domain, albeit a broad one (e.g.

a department store) whereas the dataset used in this study comprised of a variety of unrelated categories (e.g. chemistry, attire, media). Although validated by use in previous studies, any limitations would be sustained through this study. Similarly, real menus often use phrases, rather than single words. Future research could consider using a more realistic dataset that could potentially include longer labels, although care would need to be taken to ensure its academic as well as ecological validity.

Third, the TR search type lacked ecological validity – visitors rarely know the exact route to their target. This does, perhaps, strengthen the argument for mega menus, as there was a distinct preference for the mega menu when participants were unsure of the route to their target. As mentioned above, in a real-world situation it is unlikely that website visitors will know the exact route to their target and so may benefit from using a mega menu.

Fourth, as previously mentioned, three participants reported that they did not notice the sub-categories on the mega menu, so were searching the entire category dataset. Previous research has suggested that visual grouping aids search performance (Bailly et al., 2014; Brumby & Zhuang, 2015), so this may have resulted in increased search times in the mega menu conditions for these participants. To improve the current study, the delineation of sub-categories could be made more distinct. In addition, it would be interesting to control for visual design of the menus to identify if, and to what extent, these previous findings (Bailly et al., 2014; Brumby & Zhuang, 2015) are applicable to mega menus. With more links to scan, it could be expected that visual grouping is even more important in mega menus.

Fifth, whilst experiments conducted in a laboratory setting enable experimental conditions to be more easily controlled than in the real world, taking experiments out of their environmental context may effect the extent to which results can be generalised to real situations outside the laboratory. Participants were incentivised to prioritise speed and accuracy in order to simulate a real search task as much as possible. However, the formal environment and engineered circumstances of their search may have affected their performance. To mitigate these effects as much as possible, the experimenter endeavoured to put the participants at ease, by their manner and by ensuring a quiet, relaxed testing environment.

Finally, it became apparent during the experiment that some participants forgot the target item they were looking for and/or the route to the target. As a result, some were clicking several target items until they reached the correct one, instead of abandoning the trial. If the study were to be repeated, an “I’ve forgotten it” button could be included to separate these trials, or the study designed to enable participants to review the search task. However, all abandoned trials and those where the participant did not select the correct target item on the first attempt were excluded from the analysis, so this would not have adversely affected the reported results.

Design implications and future research

This study contributes to the field of human-computer interaction as it adds to the body of work on the breadth/depth trade-off by providing a first step towards understanding how mega menus perform in relation to the more established hierarchical design. Results show that mega menus do not offer performance advantages over more traditional hierarchical menus. Search type had a much greater impact on search performance than menu type. Targets with a clear path resulted in shorter search times. This suggests that aspects that strengthen the information scent and provide a clear route to the target, such as quality of link labels, are more important than menu design. Indeed, participants were much more likely to abandon trials when there was not a clear route to the target (TO searches).

Results suggest that mega menus may offer advantages in two particular situations. Firstly, mega menus may be useful on sites where the route to the target is less clear, for instance, where the size of the site makes distinct, unambiguous link labels difficult, or where lower level pages do not fit nicely into semantic groups. This is because participants showed a clear preference for the mega menu when they were not confident of where to find the target.

Secondly, mega menus may offer advantages in situations where users will have the opportunity to become familiar with the site, such as work-related systems (where users are forced to continue interacting) or sites to which users are motivated to return (such as a favourite retailer). This is because participants were less likely to abandon searches in mega menus as they became familiar with the menu.

In addition to the suggestions already made, future research could investigate the use of mega menus across particular domains (e.g. eCommerce), devices (e.g. smartphones and tablets) and user groups (e.g. the elderly, or those with physical impairments). In addition, it would be interesting to use eye tracking to identify if, and how, search strategies are affected when using a mega menu design.

Conclusion

A study was conducted which aimed to compare search performance between hierarchical menus and mega menus. The study required participants to use a hierarchical and a mega menu to complete two different search tasks: "target-only" searches where only the target item was given, and "target and route" searches where participants were given the target item and the route to get there. The results of the study suggest that items were located faster when there was a clear route to the target, and that menu design did not significantly affect search performance. Overall, participants preferred using the mega menu, especially for target-only searches where they had low confidence in where to locate the target item. Furthermore, participants were more likely to abandon target-only searches, especially when using a hierarchical menu and more likely to abandon a search in the hierarchical menu after practice.

These results are important because they present the first comparison of the traditional hierarchical menu with the more modern mega menu, and extend the breadth/depth debate to consider a menu design that offers both breadth and depth.

REFERENCES

Bailly, G., Oulasvirta, A., Brumby, D. P., & Howes, A. (2014). Model of visual search and selection time in linear menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3865-3874). ACM.

Brumby, D. P., Cox, A. L., Chung, J., & Fernandes, B. (2014). How does knowing what you are looking for change visual search behavior?. In*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3895-3898). ACM.

Brumby, D. P., & Howes, A. (2008). Strategies for guiding interactive search: An empirical investigation into the consequences of label relevance for assessment and selection. *Human–Computer Interaction*, *23*(1), 1-46.

Brumby, D. P., & Zhuang, S. (2015). Visual Grouping in Menu Interfaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4203-4206). ACM.

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* (pp. 402-409). ACM.

Card, S. K. (1982). User perceptual mechanisms in the search of computer command menus. In *Proceedings of the 1982 conference on Human factors in computing systems* (pp. 190-196). ACM.

Hochheiser, H., & Shneiderman, B. (2000). Performance benefits of simultaneous over sequential menus as task complexity increases.*International Journal of Human-Computer Interaction*, *12*(2), 173-192.

Hornof, A. J., & Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256). ACM.

Hornof, A. J., & Kieras, D. E. (1999). Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. In*Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 410-417). ACM.

Jacko, J. A., & Salvendy, G. (1996). Hierarchical menu design: Breadth, depth, and task complexity. *Perceptual and Motor skills*, *82*(3c), 1187-1201.

Katz, M. A., & Byrne, M. D. (2003). Effects of scent and breadth on use of site-specific search on e-commerce Web sites. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *10*(3), 198-220.

Kiger, J. I. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal of Man-Machine Studies*, *20*(2), 201-213.

Kreigh, R. J., Pesot, J. F., & Halcomb, C. G. (1990). An evaluation of look-ahead help fields on various types of menu hierarchies. *International Journal of Man-Machine Studies*, *32*(6), 649-661.

Krug, S. (2005). *Don't make me think: A common sense approach to web usability*. Pearson Education India.

Landauer, T. K. (1987). Relations between cognitive psychology and computer system design. *Interfacing thought: cognitive aspects of human-computer interaction*, 1-25.

Larson, K., & Czerwinski, M. (1998, January). Web page design: Implications of memory, structure and scent for information retrieval. In*Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 25-32). ACM Press/Addison-Wesley Publishing Co..

Leuthold, S., Schmutz, P., Bargas-Avila, J. A., Tuch, A. N., & Opwis, K. (2011). Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Computers in human behavior*,*27*(1), 459-472.

MacGregor, J., Lee, E., & Lam, N. (1986). Optimizing the structure of database menu indexes: A decision model of menu search. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *28*(4), 387-399.

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 and Ergonomics Society Annual Meeting*(Vol. 27, No. 9, pp. 834-837). SAGE Publications.

Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 25, No. 1, pp. 296-300). SAGE Publications.

Miller, C. S., & Remington, R. W. (2004). Modeling information navigation: Implications for information architecture. *Human-computer interaction*, *19*(3), 225-271.

Nielsen, J. (2000). Designing web usability: the practice of simplicity New Riders Publishing. *Indianapolis, Indiana*.

Norman, K. L. (1991). *The psychology of menu selection: Designing cognitive control at the human/computer interface*. Intellect Books.

Paap, K. R., & Roske-Hofstrand, R. J. (1986). The optimal number of menu options per panel. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *28*(4), 377-385.

Parkinson, S. R., Hill, M. D., Sisson, N., & Viera, C. (1988). Effects of breadth, depth and number of responses on computer menu search performance. *International Journal of Man-Machine Studies*, *28*(6), 683-692.

Parkinson, S. R., Sisson, N., & Snowberry, K. (1985). Organization of broad computer menu displays. *International journal of man-machine studies*,*23*(6), 689-697.

Pirolli, P. (2003). Exploring and finding information. *HCI Models, Theories, and Frameworks Toward Multidisciplinary Science*, *1*, 157-191.

Schultz Jr, E. E., & Curran, P. S. (1986). Menu structure and ordering of menu selection: independent or interactive effects?. *ACM SIGCHI Bulletin*,*18*(2), 69-71.

Shneiderman, S. B., & Plaisant, C. (2005). Designing the user interface 4 th edition. *ed: Pearson Addison Wesley, USA*.

Snowberry, K., Parkinson, S. R., & Sisson, N. (1983). Computer display menus. *Ergonomics*, *26*(7), 699-712.

Teitelbaum, R. C., & Granda, R. E. (1983). The effects of positional constancy on searching menus for information. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 150-153). ACM..

Tullis, T. S. (1985). Designing a menu-based interface to an operating system. In *ACM SIGCHI Bulletin* (Vol. 16, No. 4, pp. 79-84). ACM.

Vandierendonck, A., Van Hoe, R., & De Soete, G. (1988). Menu search as a function of menu organization, categorization and experience. *Acta Psychologica*, *69*(3), 231-248.

Yoon, C., Feinberg, F., Hu, P., Gutchess, A. H., Hedden, T., Chen, H. Y. M., & Park, D. C. (2004). Category norms as a function of culture and age: comparisons of item responses to 105 categories by american and chinese adults. *Psychology and aging*, *19*(3), 379.

Zaphiris, P. G. (2000). Depth vs breath in the arrangement of web links. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, No. 4, pp. 453-456). SAGE Publications.

Zaphiris, P. (2001). Age differences and the depth-breadth tradeoff in hierarchical online information systems. In HCI (pp. 540-544).

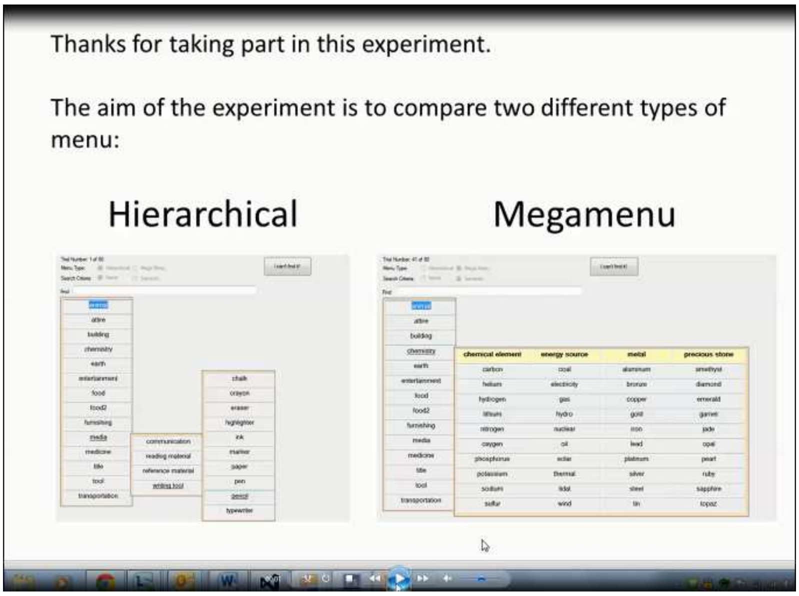
Zaphiris, P., Shneiderman, B., & Norman, K. L. (2002). Expandable indexes vs. sequential menus for searching hierarchies on the World Wide Web.Behaviour & Information Technology, 21(3), 201-207.

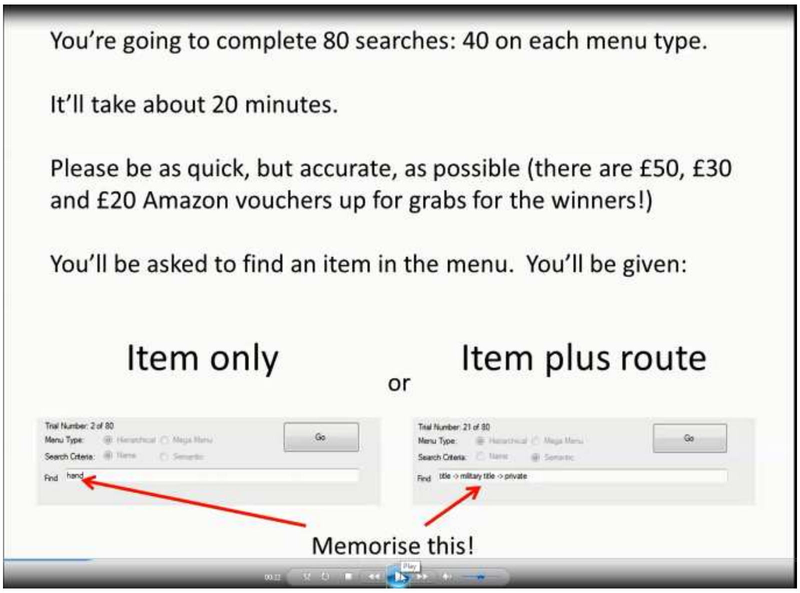
APPENDICES

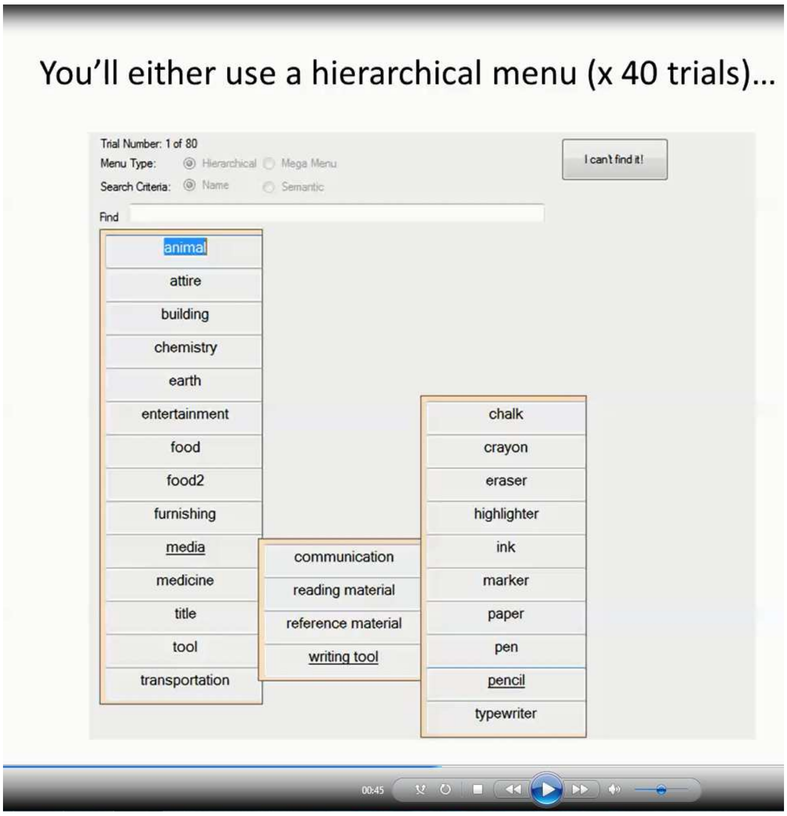
Appendix A - Semi-structured interview questions

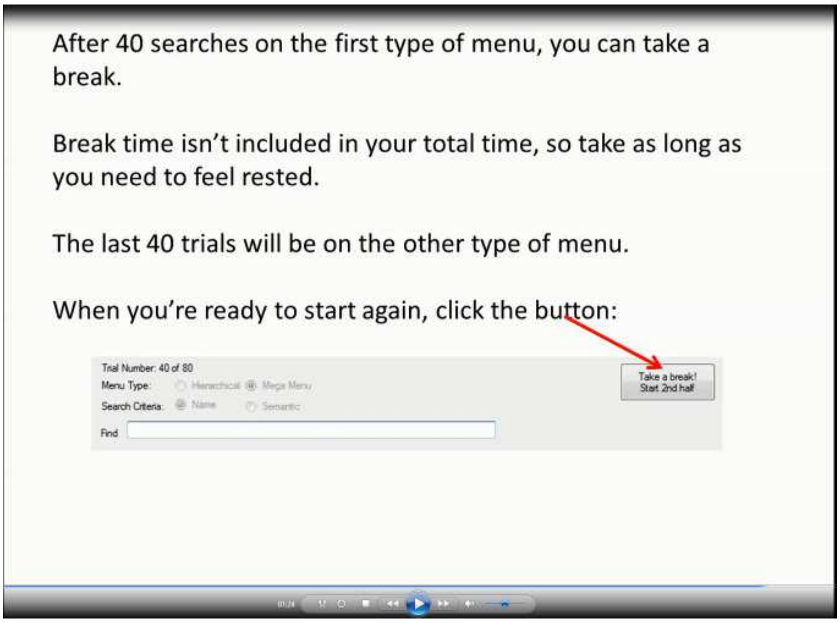
1. How did you find that?
2. Which menu did you prefer?
3. What was it about that menu that you liked/disliked?
4. Which menu did you prefer when you were given the target and the route to get there?
5. What about when you were just given the target word?
6. What about the menu – what did you like/dislike about it?
7. Which search task did you prefer – when you given just the target or also the route to get there?
8. Any other thoughts?

Appendix B - Instructional video (screenshots)






You'll either use a hierarchical menu (x 40 trials)...
Trial Number: 1 of 80
Menu Type: Hierarchical Mega Menu
Search Criteria: Name Semantic
Find
I can't find it!
animal
attire
building
chemistry
earth
entertainment
food
food2
furnishing
media
medicine
title
tool
transportation
communication
reading material
reference material
writing tool
chalk
crayon
eraser
highlighter
ink
marker
paper
pen
pencil
typewriter
00:45


After 40 searches on the first type of menu, you can take a break.
Break time isn't included in your total time, so take as long as you need to feel rested.
The last 40 trials will be on the other type of menu.
When you're ready to start again, click the button:
Trial Number: 40 of 80
Menu Type: Hierarchical Mega Menu
Search Criteria: Name Semantic
Find
Take a break! Start 2nd half

Appendix C - Instruction sheet and consent forms

Template Information Sheet for Participants in Research Studies

You will be given a copy of this information sheet.

Title of Project: **Comparison of hierarchical and megamenus**

This study has been approved by the UCL Research Ethics Committee as Project ID Number: UCLIC/1415/007/MSc Brumby/Naylor

Name, Address and Contact Details of Investigators:

Dr Duncan Brumby, d.brumby@ucl.ac.uk

Samantha Naylor Samantha.naylor.11@ucl.ac.uk

University College London Interaction Centre, Malet Place Engineering Building (8th Floor), UCL, Gower Street, London, WC1E 6BT.

We would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, please read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or you would like more information.

Insert Details of Study

The purpose of this research is to compare searching for target words in two types of menu: hierarchical and megamenus. Two types of search will be undertaken: one will simply give the target word, and the other will give the target word and the route through the navigation. You will be asked to complete 40 trials on each type of menu (with a short break in between). If you can't find a target, you can abandon the search by clicking the "I can't find it!" button.

Please be as quick and accurate as possible. At the end, you will be asked for your preferences between the two types of menu – there are no right or wrong answers to this. No risks or discomfort is expected. The session will last approximately 30 mins in total.

It is up to you to decide whether or not to take part. If you choose not to participate, you won't incur any penalties or lose any benefits to which you might have been entitled. However, if you do decide to take part, you will be given this information sheet to keep and asked to sign a consent form. Even after agreeing to take part, you can still withdraw at any time and without giving a reason.

All data will be collected and stored in accordance with the Data Protection Act 1998.

Informed Consent Form for Participants in Research Studies

(This form is to be completed independently by the participant after reading the Information Sheet and/or having listened to an explanation about the research.)

Title of Project: **Comparison of hierarchical and megamenus**

This study has been approved by the UCL Research Ethics Committee as Project ID Number: UCLIC/1415/007/MSc Brumby/Naylor

(Ctrl)

Participant's Statement

I ……………………………………………………………………………

agree that I have

- read the information sheet and/or the project has been explained to me orally;
- had the opportunity to ask questions and discuss the study; and
- received satisfactory answers to all my questions or have been advised of an individual to contact for answers to pertinent questions about the research and my rights as a participant and whom to contact in the event of a research-related injury.

I understand that I am free to withdraw from the study without penalty if I so wish, and I consent to the processing of my personal information for the purposes of this study only and that it will not be used for any other purpose. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Signed: Date:

Investigator's Statement

I ……………………………………………………………………………

confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signed: Date:

- I understand that my system usage will be recorded, and will be used for researcher purposes only – no recordings will be published
- I understand that the information I have submitted will be published as a report. Confidentiality and anonymity will be maintained, and it will not be possible to identify me from any publications.

Appendix D – Data captured

Field name	**Data collected**
partID	Participant ID (A – P)
Counter	Trial ID (1 – 80)
menuType	Either “hierarchical” or “mega menu”
Abandoned	Whether the participant abandoned the search by clicking the “I can’t find it!” button
searchTextType	Either “Target only” or “Route and target”
blockID	1 – 2, showing whether the trial was in the first block of 40 trials or the second block of 40 trials
blockTrialId	1 – 10, showing each trial within each block of 10 trials
myTime	Time taken from clicking “New search” to locating the target item
pauseTime	Time taken between locating the target item (upon which the next search description would be displayed) and starting the next search (by clicking the “Go” button)
breakTime	Time taken between finishing the first block of 40 trials and starting the second block of 40 trials
myErrors	Number of times an incorrect target was clicked
myCategory	Category of the target item
mySubCategory	Sub-category of the target item
menuItem	Name of target item
indexOfCategory	1 – 14 showing position in category list
indexOfSubCategory	1 – 4 showing position in sub-category list
indexOfMenuItem	1 – 10 showing position in menu item list