THE EFFECT OF SCENT ON USER RECALL OF WEB SITE NAVIGATION

Marc L. Resnick and Amir Abou Baker
Industrial and Systems Engineering
Florida International University
Miami, FL

When designing navigation menus there are some key factors usability professionals struggle with to ensure an easy to use navigation system. In the short term, the menu should support effective user wayfinding and confidence to accomplish the current task. Menus can also achieve longer term benefits by intimating the information architecture of the site, providing support for future tasks. The effectiveness of menu design hinges on whether these objectives are achieved. The focus of this study was the impact of label scent on user performance in navigation tasks. Labels of high and low scent were placed at two menu positions to measure the impact on the short term goal of finding a known item and the long term goal of developing an awareness of the information architecture. The results provide insights that can be used in the design of navigation menus.

INTRODUCTION

In the design of navigation menus, there are two objectives that often compete. On one hand, we would like to design the menu so that users find the content they are seeking quickly and confidently. Spool (2003) reported that high scent menu items (words that match the users' goals in their own terminology) achieve both speed and user confidence. Katz and Byrne (2003), who defined scent as the overlap between the menu item and the user's mental model of their objective, also found higher task success when a high scent item is located in the navigation menu. Resnick and Sanchez (2004) found this benefit at several levels of scent strength. They found linear relationships between menu label scent and both navigation errors and navigation time. This is particularly important on gallery pages, which are the most critical link to content (Spool, 2005).

The position of the target on the navigation menu is also important. Users most often apply a self-terminating threshold strategy (Miller et al., 2004; Weinreich et al., 2006), selecting the first item on the menu that has a sufficient scent rather than reading them all and selecting the one with the highest scent (an exhaustive comparison strategy). Therefore, items at the top of the menu will be identified and selected faster than those further down the menu. Ideally, the target item will have a high scent and be located at the top of the menu. This combination would maximize user confidence and minimize navigation time and navigation errors.

The consequence of this design is that users find their targets without getting any broader sense of the site's contents or organization. A secondary goal of information architecture is to support the user in developing a mental model of the site's content to support future task performance. A further benefit of an accurate mental model is the enhanced opportunity for cross-selling additional products and services. This will not happen when users find their target too efficiently. Oulasvirta et al. (2005) found that menu items that are not considered task-relevant are not recalled. It seems that any memory of menu items that are merely skimmed is fleeting at best. Ironically, low scent menus force users to process each item more deeply to make their navigation decision and thus may enhance the development of broader mental models of the site's content. While intentionally using low scent menu items would not be a successful Information Architecture strategy, the specific effects on user performance warrant further investigation. A quantitative understanding of the tradeoffs required in each design decision will enable better designs overall.

METHODS

Experimental Design

Two variables were tested in this study. The first variable was the position of the target menu item. The target item was located either at the first or fifth position in a ten-item menu. The second variable was the scent of the target menu item label. This target item was either a high scent (strong association with the target item) or low scent (weak association with the target item) label for the assigned user task. Terms were selected from those used in Resnick and Sanchez (2004). That study involved extensive pre-testing to measure the scent of labels for a set of specific e-commerce tasks.

Table 1: Design Variables

Variable	Levels
Target Label Position	First, Fifth
Target Label Scent	High, Low

The two variables were then crossed in a 2x2 factorial pairing to measure the main effects of each variable and the interactions between them. The result was four unique test conditions that were examined:

- First position, high scent
- First position, low scent
- Fifth position, high scent
- Fifth position, low scent

Test Materials

Four versions of the gallery page of a nutritional e-commerce website were created based on the experimental conditions listed in Table 1. All other aspects of the page remained the same with no change. A sample screen shot of one of the gallery pages is presented below in Figure 1.

Test Participants

One hundred participants were recruited as test subjects for the study. Twenty five participants were assigned to each experimental condition. The participants were randomly selected college students and professionals. There were 51 female and 49 male participants ranging in age from eighteen to thirty-five.

Apparatus

A HP Compaq 6910p Laptop with 15.4” widescreen display set to a 1280 x 800 resolution and using a standard laptop keyboard and mouse was used to run all tests.

Test Procedure

A scenario was developed to simulate a naturalistic environment and a realistic set of products. In the scenario the participant was asked to visit the site then was assigned the task of finding a certain product for sale. Each participant was given the same scenario and was assigned the same task. The user was instructed to find a specific item for sale on the site by using the navigation menu.

When the test participant selected an item from the menu, he/she was taken directly to a data collection interface and asked to rate his/her confidence in the menu item selection using the pre-click confidence scale (Resnick and Lergier, 2003). This provided a quick and user-specific measure of the scent of the item.

A series of recall questions were then administered. The first question asked the user to recall the precise term that he/she had selected from the menu using a free recall format. This test was used to estimate the depth of processing of the menu item that the participant selected during the task. When this was completed, the user was asked to identify the rest of the items from the navigation menu in a recognition format. Ten semantically-related items were presented to the participant. Five of these items were actual items from the menu while the remaining five were distracters. The participant was asked to select the items from the list he/she thought were present on the navigation menu. Recall for items that were included among the original menu items were compared to the recall of distracter items that were not located in the menu. This was used to assess each participant's mental model of the site's overall information architecture.

Performance Metrics

Several metrics were used to evaluate the participant's performance on the navigation task and the development of a mental model of the information architecture.

Navigation Task:

1. Selection accuracy: Did the user select the correct item from the menu?
2. Selection speed: How much time did it take for the user to select a menu item?
3. Selection confidence: How confident was the user that he/she selected the correct menu item? This was measured using the a five-point Likert scale.
4. Selection recall: Did the user recall the term that he/she selected from the menu?

Information Architecture

5. Recall of target menu items: How many target items from the menu did the user recall? This was measured relative to the number of distracters also selected to account for detection bias.
6. Recall of non-target menu items: How many non-target items from the menu does the user recall? This was also measured relative to the number of distracters selected to account for detection bias.

RESULTS

Task Performance

A binary logistic regression was used to evaluate the effects of scent and target item position on selection accuracy. Scent had a significant main effect on accuracy ($p = 0.048$). When the target item was a low scent label, only 42% of participants accurately selected it, compared to 86% when the target item was a high scent label (see Figure 2). There was no effect of target item position, participant gender, or any interactions. Analysis of Variance was used to evaluate the effects of scent and target item position on the selection speed and participant pre-click confidence. No significant effects were identified for speed. Scent had a marginally significant effect on confidence ($p = 0.089$) but an ordinal logistic regression failed to confirm this result. All participants in all conditions were able to successfully recall the term in the menu that they selected in the free recall task.

Mental Model Development

Analysis of Variance was used to evaluate the effects of target item scent and target item position on the number of menu items that could be identified in the recognition test. There was a significant effect of scent ($p = 0.000$). After completing the task with a low scent target, 42 participants identified three or more items from the navigation menu with a mean of 3.5. After completing the task with a high scent target, only 27 participants identified three or more items from the navigation menu with a mean of 2.5 (see Figure 3). There was a marginally significant effect of target item position ($p = 0.078$). When the target item was in the first position, the mean number of items identified from the navigation menu

was 2.8 compared to 3.2 when the target item was in the fifth position. Similar results were found for the number of distracters incorrectly identified as menu items.

DISCUSSION

There are implications of these results for both the short term objective of supporting effective navigation and wayfinding and the long term objective of helping uses to develop mental models of a site's information architecture. This study confirms previous results that scent has a significant impact on successful navigation. High scent category labels allowed twice as many participants to identify the correct category on their first attempt. This is a critical factor in minimizing user frustration and generating customer loyalty. The advantage of high scent labels on pre-click confidence did not reach statistical significance but taken in combination with previous studies suggests that scent does imbue confidence. Confident users are also more likely to be satisfied and loyal.

It was expected that the task time variable would provide insights into whether participants used self-terminating or exhaustive strategies when scanning navigation menus. However, these results did not reach significance. The raw data suggest that many different strategies were used that mask the overall results. Some participants seemed to scan the menu starting from the bottom. Eye-tracking studies may be needed to resolve these questions. The finding that more non-target menu items were recalled after completing the task with the target at the fifth position than the first position suggests that most participants did start from the top and used a self-terminating strategy. However the marginal significance of this result requires confirmation from other studies.

The primary focus of this study was on the longer term objective of helping users to develop mental models of a site's information architecture and the tradeoffs that may result from high scent navigation labels. This is where the results show the difficult challenges of navigation design. All users were able to recall the precise term that they selected from the navigation menu. There was only a few second delay before the recall test was administered. This may have reduced the sensitivity of the approach. A longer delay may show that lower scent items are more deeply processed at the lexical level because of the semantic uncertainty.

The effect on recall for the rest of the menu items was significant and supports the main hypothesis of the study. After completing the task with a low scent target, participants recalled significantly more non-target menu items. This suggests that those other items were more deeply processed, although the study did not address whether this occurred at the semantic or lexical level. This extra processing can lead to better mental models of the overall site's information architecture. It is unclear how this was achieved without an increase in task time.

But this result does present a tradeoff between the short term and long term objectives of menu design. High scent labels are clearly superior for finding specific items. But it comes at the cost that users do not need to process other menu items and thus do not get a sense of what else is available for future tasks.

CONCLUSION

This research was intended to provide insights into the effects of high and low scent labels on short term and long term user performance. These insights can be used by information architects and menu designers to achieve the specific needs of individual systems. However, the study only begins to investigate the relationship between short term wayfinding success and long term mental model development. Future research is necessary to identify the combination of design attributes that most effectively balances short term task success and long term familiarity, satisfaction, and loyalty to a system.

These results can serve as the basis for a more extensive that looks at multiple menu positions and uses eye-tracking to better resolve menu scanning patterns and strategies. Such an extensive study would

provide practitioners with cost/benefit data to guide the design of a gallery of navigation structures and designs.

REFERENCES

Katz M.A. and 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*, 10, 3, 198-220.

Miller C.S. and Remington R.W. (2004). Modeling information navigation: implications for information architecture. *Human-Computer Interaction*, 19, 225-271.

Oulasvirta A., Karkkainen L. and Laarni J. (2005). Expectations and memory in link search. *Computers in Human Behavior*, 21, 773-789.

Resnick M.L. and Lergier R. (2003). Task specific user strategies in on-line search. *Journal of E-Business*, 3, 1, 1-22.

Resnick M.L. and Sanchez J. (2004) The effects of organizational scheme and label quality on task performance in product-centered and user-centered retail web sites.*Human Factors,* 46, 1, 104-125.

Spool J.M. (2003). Getting confidence from lincoln. *UIETips*, Retrieved on 4/28/03 at www.uie.com/articles/getting_confidence.htm.

Spool J.M. (2005). The eight types of navigation pages. *UIETips,* Retrieved 12/01/05 at www.uie.com/brainsparks/2005/11/28/the-8-types-of-navigation-pages.

Weinreich H., Obendorf H., Herder E. and Mayer M. (2006). Off the beaten tracks: exploring three aspects of web navigation. *Proceedings of WWW 2006*. Association for Computing Machinery: New York, NY.

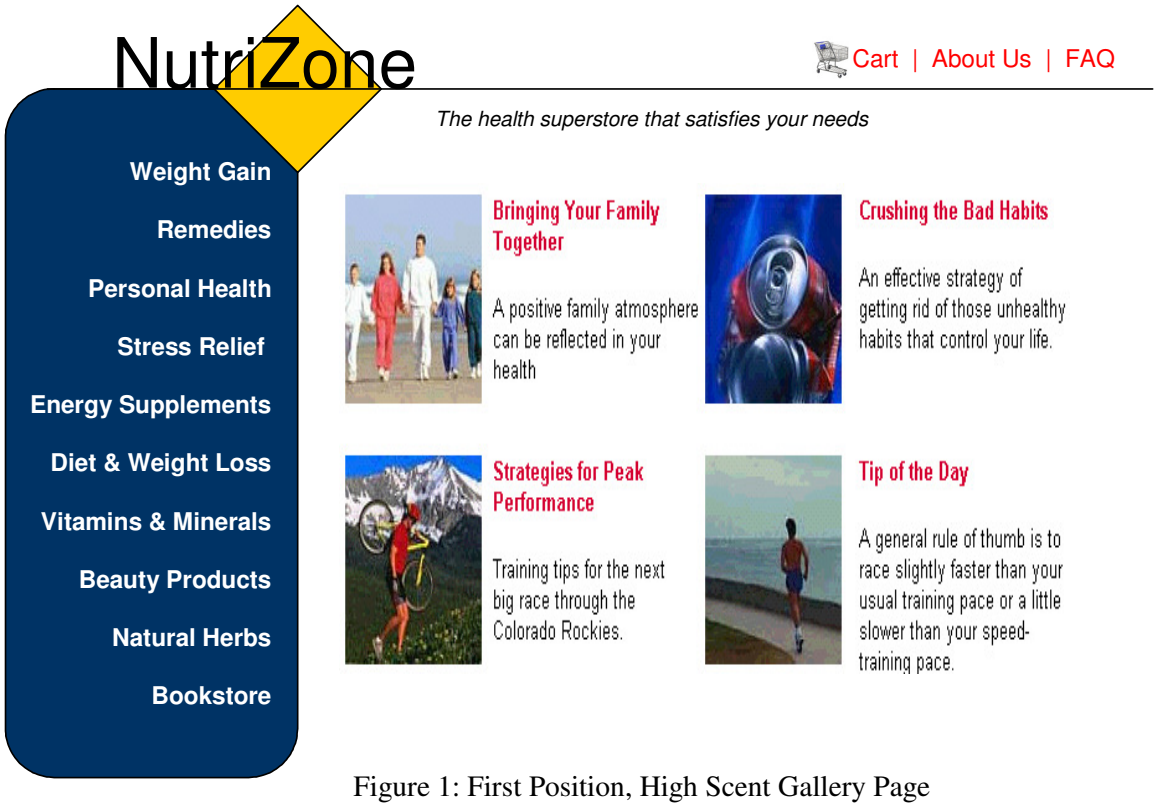

Figure 1: First Position, High Scent Gallery Page

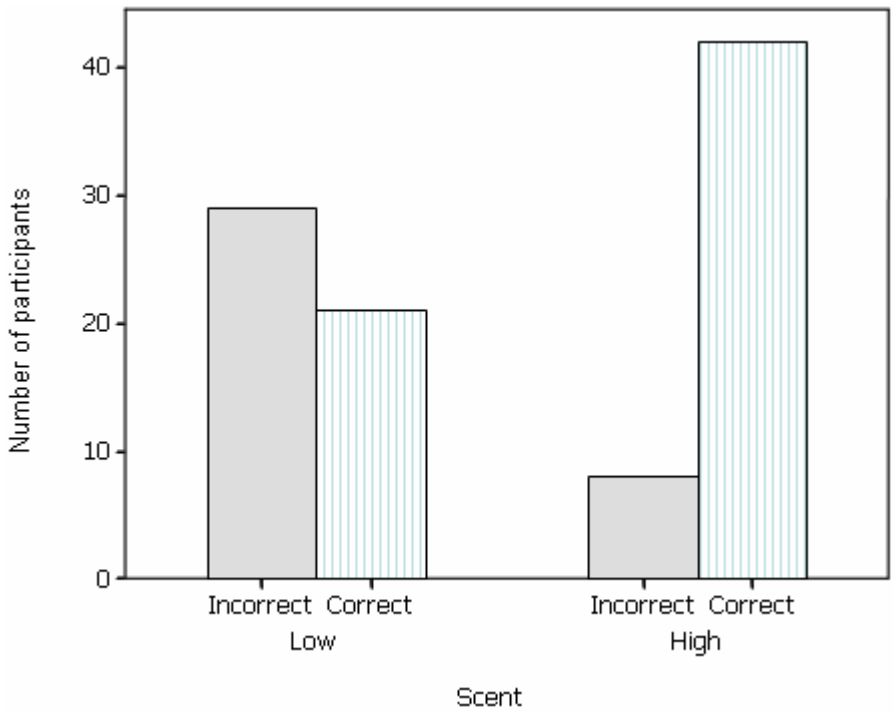

Figure 2. Number of participants correctly selecting the target menu item in low and high scent conditions

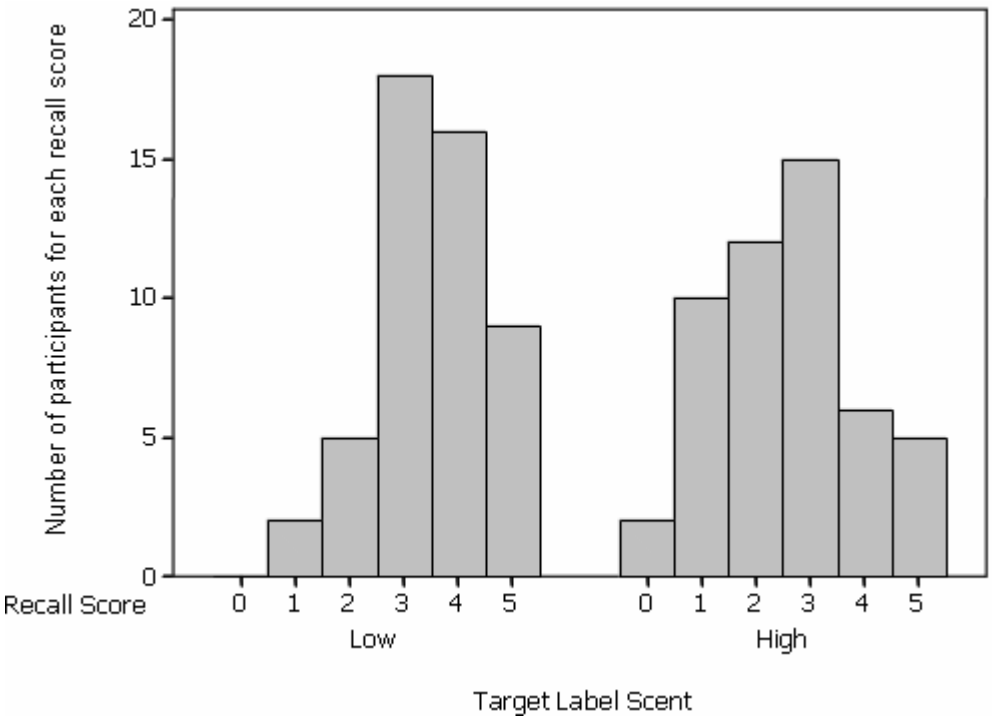

Figure 3. Number of non-target menu items correctly recalled from the navigation menu for low and high scent conditions