# Evaluating the Cognitive Impact of Search User Interface Design Decisions

**Max L. Wilson**
Future Interaction Technology Labs
Department of Computer Science, College of Science
Swansea University, UK
m.l.wilson@swansea.ac.uk

### ABSTRACT
The design of search user interfaces has developed dramatically over the years, from simple keyword search systems to complex combinations of faceted filters and sorting mechanisms. These complicated interactions can provide the searcher with a lot of power and control, but at what cost? Our own work has seen users experience a sharp learning curve with faceted browsers, even before they begin interacting. This paper describes a forthcoming period of work that intends to investigate the cognitive impact of incrementally adding features to search user interfaces. We intend to produce search user interface design recommendations to help designers maximize support for searchers while minimizing cognitive impact.

### Author Keywords
Search, Exploratory Search, User Interface Design, Cognitive Load Theory

### ACM Classification Keywords
H5.2. Information interfaces and presentation (User Interfaces): evaluation/methodology, screen design. H3.3. Information search and retrieval: Search process.

### INTRODUCTION
User Interface (UI) Designers are always concerned with supporting users effectively and intuitively, but a common recent focus for Search User Interface (SUI) designers has been to increase the interactive power and control that searchers have over results. As a community, we want to support users in exploring, discovering, comparing, and choosing results that meet their needs. SUI designers, therefore, are concerned with maximizing the use of powerful interface features while maintaining a clear and intuitive design.

In our prior work, we developed mSpace [7] as a faceted browser that lets searchers use combinations of orthogonal metadata filters to narrow their search. We developed advanced interactions for faceted browsers that took advantage of visual location within the SUI, and highlighted options in unused filters that were related to guide searchers [10]. Frequently, however, we informally noted that searchers spent increasing periods of time on visually comprehending the interface before making their first move. In follow up studies, we saw minimal interaction with facets during the first visit, but recorded a significant increase in the use of faceted features during subsequent return visits. It is the hypothesis of our forthcoming work that this non-use of such powerful features is caused by an increased cognitive load created by the associated increased complexity of the SUI. It is this cognitive impact that we believe can be measured and attributed to specific design decisions.

mSpace is one specific faceted browser, but the principle of faceted browsing can be implemented in many different ways [2]. We also hypothesize that not only the presence, but also the subsequent design of SUI features can also have an impact. The following sections cover some related work before describing our plans to evaluate the cognitive impact that adding features to SUIs can have.

### RELATED WORK
SUI design is affected by many factors. Interaction designers can decide how best to support searchers, but designs may be limited by the metadata that is available about the possible results. Both the underlying data and the graphical design may also have an impact, then, on how the chosen interaction will look and feel. As perhaps the most recognized SUI for many users around the world, Google has always maintained a very clean and clear white design[1], and make very incremental careful design changes that stay within that design. Competitor search engines have notably changed over the years, with many now being very similar to Google in terms of interaction design, while trying to keep their own visual design consistent.

For more exploratory websites that sell a wide range of products, or provide large collections of information or documents, there are now many different features that support people, from tabular or dropdown-based sorting



[1] http://searchengineland.com/qa-with-marissa-mayer-google-vp-search-products-user-experience-10370

mechanisms, to categories, clusters, filters, and facets. Some websites that provide these features are frustrating and difficult to use, while others are simple, intuitive, and successful. In these systems it is often the way that the ideal support has been developed that has affected their success. In a study of the success of different faceted browser implementations, Capra et al [1] directly compared two faceted browsers to a government website, all over the same hierarchical government dataset, and discovered that the customized hierarchical design of the original website supported searchers far better than the functionally more powerful faceted browsers.

Both the choice of content and the visual design have both been shown to have an impact on usability. White et al showed that the text that includes the search terms is best, and that highlighting these terms also improves search [12]. Similarly, Lin et al. have shown that simply highlighting the domain name in the URL bar significantly reduces the chances that users will be caught be fishing attacks [4]. Zheng et al [13] have also shown that users can make often-accurate snap judgments about the credibility of websites within half a second. Further, Wilson et al [10] noted that the success of adding guiding highlights to their faceted browser was affected by the choice of highlight-colour and its implied meaning.

The choice of SUI features within a single implementation has also been shown to have an impact on search success. Diriye et al compared a keyword search interface with a revised version that also included query suggestions [3]. Their results showed that such features slowed down searchers who were performing simple lookup tasks, but supported those who were performing more complicated exploratory tasks. Similarly, Wilson and Wilson have also found early results indicating that the simple presence, without interaction, of a keyword cloud provides additional support, where subsequent interaction provides very little gain [11] during exploratory tasks. Wilson and Wilson's results suggest that searchers can learn more about the result set from seeing the terms in the keyword cloud, than actually using them to filter the results.

The location of features within a SUI has also been shown to have an impact. Morgan and Wilson studied the visual layout of search thumbnails, predicting that having a rack of thumbnails at the top of the user interface would allow searchers to make faster judgments when trying to re-find pages [5]. Their results showed that a rack of thumbnails was significantly more disruptive to searchers when the target page was not in the results, than the support it provided when it was.

The studies above indicate that the success of SUIs can be attributed to the appropriateness of the functionality provided, where unnecessary functionality can slow users down. Further, the studies indicate that the success of SUIs can be determined by simple visual or spatial changes that do not necessarily impact functionality. Consequently, where two systems provide the same support, one may be harder or easier to use because of its simple visual design. Our conclusion is that to understand the success of a SUI, we must analyse both the support in terms of functionality, and the cognitive impact is creates. Being able to understand and predict these two things would help us to design and build better SUIs

## EVALUATING THE SUPPORT PROVIDED BY SUIS

Beyond the common practice of performing task-oriented user studies, my own doctoral work focused on the design of an analytical evaluation metric for SUIs, called the Search Interface Inspector[2] (Sii). Sii calculates the support for different types of users based upon the set of features in the interface, and how many interactions they take to use [9]. To analyse a SUI, the evaluator catalogues the features of the design and calculates how many interactions are required to perform a set of known search tactics. The method then interpolates the likely support for different types of searchers (explorers or searchers that know what they are looking for, for example), based upon the types of tactics they are likely to perform. Sii can be used to compare several designs and produces a series of 3 interactive graphs that allow evaluators perform an investigative analysis of the results.

Sii is based on detailed established information seeking theory and rewards the design of search functionality that has simple interaction. Consequently, however, Sii rewards the addition of new simple functionality, without being able to estimate the increasing complexity of the SUI as new features are added. To remedy this problem, a chapter of the thesis investigated Cognitive Load Theory and initially specified a similar metric that calculated the cognitive load of a UI. This second measure of intrinsic cognitive load was proposed for inclusion in Sii, estimated the intrinsic cognitive load of a SUI. Similar to how the original metric was correlated with study results, one aim of the work described below is to further refine and validate this analytical measure of the cognitive impact of SUIs.

Cognitive Load Theory highlights that capacity for learning is affected by three aspects: intrinsic, extrinsic, and germane cognitive load. Intrinsic cognitive load is created by the materials providing the learning experience, or in our case the SUI. Extrinsic cognitive load is created by the complexity in the task at hand. Germane cognitive load is then required to process what is learned and commit it to long-term memory. If intrinsic load and extrinsic load are too high, then there may not be enough space load left for germane cognitive load. Although, it is commonly accepted that effort can increase overall capacity, the aim should still be to reduce intrinsic cognitive load by improving the design of learning materials or SUIs [6]. Reducing intrinsic load creates space for users to perform increasingly

[2] http://mspace.fm/sii

complex tasks, or opens-up germane cognitive load so that what is being learned can be retained.

## EVALUATING THE COGNITIVE IMPACT OF SUIS

The general structure of the studies we are planning is to use brain scanners to record the cognitive impact that different SUIs have on a user. The initial phases will focus on identifying and measuring such responses to significant and obvious differences, before trying to capture changes to more subtle designs and, hopefully, in-situ. Initially, we will be using EPOC Emotiv headsets[3], as shown in Figure 1, to take readings. These headsets are commercialized versions of EEG scanners, but are designed for use in more natural contexts. EEG scanners, as with many other brain scanning systems, are typically affected by simple body movements and so are often restricted to confined conditions. Such scanners, therefore, are often not suitable for task-based evaluations, which require action and movement. In psychology, EEG scanners are typically used in constrained environments where users are only allowed to move their thumbs to answer yes or no. Consequently, this work requires scanners that can be used in more natural contexts while performing everyday searching tasks. In the future, funding permitting, we also intend to buy an fNIR scanner, which has been shown to be suitable for task-based evaluation conditions [8]. We intend to use these measurements to understand the impact of design decisions, in order to make clear recommendations to SUI designers.

**Figure 1: EPOC Emotiv Headset**

### Phase 1 – the impact of additional features

Beginning this summer, with two summer interns, we will be performing our first studies, which will simply display SUIs of incremental complexity to participants. We will begin with a simple keyword search design, and add features such as recommendations and filters. The order that interfaces are shown to participants will be randomized to avoid learning and familiarity bias. The aim of this phase is to prove that the learning curves experienced by users exist and the cognitive load can be measured objectively. We hope that the results will show initial insight into the amount of impact that different features have, which may in turn help us make hypotheses about design issues. This phase will help us identify the cost of adding a feature, where task success would allow us to measure their benefit.

### Phase 2 – capturing impact in the context of tasks

Where the first phase above allows us to learn to recognize the signs from EEG signals, we intend to try and detect cognitive load in situ, and in the context of a task. We will be setting participants specific simple and exploratory tasks, whilst controlling the type of user interface features they see, to capture the cognitive impact as they start. This phase will help us identify whether the impact of a search user interface is affected by task context.

### Phase 3 – the impact of different implementations

While adding features creates an obvious change in the user interface, different features can be put in different places in the SUI and also be implemented differently. Google, for example, puts suggested refinements at the bottom of the page, while Bing has them on the side. Bing also chooses to provide a mix of refinements and alternative directions. In Phase 2 we intend to analyse both of these kinds of variables to see if they have significant impacts on cognitive load. This phase will help us identify whether the cost of adding SUI features can be minimized by refining their design.

### Discussion

There are many challenges remaining in this planned work. So far, we have planned very controlled comparisons of SUI changes, but in real life these systems are used in the context of complex tasks and for extended periods of time. Controlled situations will help identify cause and effect, but other similar objective measurements, like eye trackers, still require interpretation. We hope to expand on these methods, and the findings of existing brain scanning HCI research [8], by addressing this issue over time. Finally, although this research is primarily interested in the development of SUI interfaces and how they affect people learning to use powerful search features, there are many other things that can be distracting in general UI design. These methods will likely expand to help address other design questions; we, however, are particularly aiming to answer questions about encouraging exploratory search and learning, by increasing the power of SUIs, while reducing their impact on searchers.

## CONCLUSIONS

This work has yet to begin formally, but we intend to learn more about the impact that very simple design decisions can have on searchers. From previous experience of searcher success in evaluations, both industry and academia know that such changes can seriously impact the success of a search user interface. This work will use objective measurements of brain response to help us identify the factors that make search user interfaces hard to comprehend. We hope that such measurements will a) help us analyse the cost-benefit trade-off of adding additional

---

[3] http://www.emotiv.com/

support to search user interfaces, and b) help us develop design recommendations for implementing search user interface features so that they have minimal impact.

## REFERENCES


1. Robert Capra, Gary Marchionini, Jung Sun Oh, Fred Stutzman, and Yan Zhang. 2007. Effects of structure and interaction style on distinct search tasks. In *Proc. JCDL '07*. ACM, New York, NY, USA, 442-451.
2. Edward C. Clarkson, Shamkant B. Navathe, and James D. Foley. 2009. Generalized formal models for faceted user interfaces. In *Proc. JCDL '09*. ACM, New York, NY, USA, 125-134.
3. Abdigani Diriye, Ann Blandford, and Anastasios Tombros. 2010. Exploring the impact of search interface features on search tasks. In *Proc. ECDL'10*.
4. Eric Lin, Saul Greenberg, Eileah Trotter, David Ma, John Aycock. Does Domain Highlighting Help People Identify Phishing Sites. In *Proc. CHI2011* (in press).
5. Rhys Morgan and Max L. Wilson. 2010. The Revisit Rack: grouping web search thumbnails for optimal visual recognition. In *Proc. ASIS&T '10*.
6. Sharon Oviatt. 2006. Human-centered design meets cognitive load theory: designing interfaces that help people think. In *Proc. MULTIMEDIA'06*. ACM, New York, NY, USA, 871-880.
7. m.c. schraefel, Max Wilson, Alistair Russell, and Daniel A. Smith. 2006. mSpace: improving information access to multimedia domains with multimodal exploratory search. *Commun. ACM* 49, 4 (April 2006), 47-49.
8. Erin Treacy Solovey, Audrey Girouard, Krysta Chauncey, Leanne M. Hirshfield, Angelo Sassaroli, Feng Zheng, Sergio Fantini, and Robert J.K. Jacob. 2009. Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. In *Proc. UIST '09*. ACM, New York, NY, USA, 157-166.
9. Max L. Wilson, M. C. schraefel, and Ryen W. White. 2009. Evaluating advanced search interfaces using established information-seeking models. *J. Am. Soc. Inf. Sci. Technol.* 60, 7 (July 2009), 1407-1422.
10. Max L. Wilson, Paul André, and mc schraefel. 2008. Backward highlighting: enhancing faceted search. In *Proc UIST '08*. ACM, New York, NY, USA, 235-238
11. Wilson, M. J. and Wilson, M. L. Tag Clouds and Keyword Clouds: evaluating zero-interaction benefits. In *Ext. Abstract CHI'11.*
12. Ryen W. White, Ian Ruthven, and Joemon M. Jose. 2002. Finding relevant documents using top ranking sentences: an evaluation of two alternative schemes. In *Proc. SIGIR '02*. ACM, New York, NY, USA, 57-64.
13. Xianjun Sam Zheng, Ishani Chakraborty, James Jeng-Weei Lin, and Robert Rauschenberger. 2009. Correlating low-level image statistics with users - rapid aesthetic and affective judgments of web pages. In *Proc. CHI '09*. ACM, New York, NY, USA, 1-10.