

An Inquiry into Mental Models of Web Interface Design

Completed Research Paper

Daniel S. Soper and Sinjini Mitra

Department of Information Systems and Decision Sciences

Mihaylo College of Business and Economics

California State University, Fullerton

dsoper@fullerton.edu • smitra@fullerton.edu

ABSTRACT

Here we adopt a cognitive science orientation to inquire into web interface design using the lens of mental models. A mental model is an internal cognitive representation of some aspect of the real world. The human mind automatically constructs and refines these models to provide a basis from which we can reason, make decisions, and form expectations. Even without conscious awareness, cognitive theory suggests that each web user has developed a mental model of what constitutes agreeable web interface design. But what is the nature of these mental models? Further, to what extent do people possess a general, shared mental model of web interface design? We examine these questions using a novel methodological approach in conjunction with statistical analyses and thermal imaging. We find not only that people possess a shared mental model of web interface design, but also that this shared model is remarkably cohesive from person to person.

Keywords

Web interface design, mental models, blank screen method

INTRODUCTION

From the perspective of usability, all websites are certainly not created equal. Indeed, most web users occasionally encounter a website which seems poorly designed or difficult to use. Navigation may be challenging, advertisements may be overly obtrusive, or perhaps certain content or features may be difficult to locate. Regardless of the specific reasons for our discomfort – and regardless of whether we are consciously aware of these reasons – something simply feels “wrong” about the design of certain websites. The discomfort that we feel at these times can have important consequences for organizations that publish and maintain such websites (Palmer, 2002). Users can become frustrated (Nadkarni and Gupta, 2007), become disloyal and lose faith in the publisher (Flavián et al., 2006), or find their opinions of the underlying organization declining (Agarwal and Venkatesh, 2001). As a result, organizations can lose not only customers and their attendant revenues, but can also suffer losses in reputation and customer goodwill. Every organization which maintains a web presence should therefore be highly interested in ensuring that the design of its website is aligned with the expectations of web users. Unfortunately – as most of us can attest from personal experience – this is often not the case.

The question of why certain websites seem poorly designed is a daunting one, and answering that question in its entirety requires a multifaceted approach. To begin, one must define precisely what is meant by “website design”, as the phrase itself is imbued with ambiguity. Depending upon the context, “website design” might alternately refer to a web designer’s choice of colors or fonts, the ease with which the website can be navigated, or the layout of the web interface. Research seeking to identify what constitutes appropriate or agreeable website design must therefore carefully define its scope of inquiry. In that spirit – and acknowledging that we certainly do not dismiss the relevance of typography, color, or navigability – our efforts here are constrained to the way in which the various elements that together comprise a web interface should be properly arranged in order to meet user expectations. For this purpose we draw heavily on research into human cognition, and place at the core of our inquiry a key theoretical concept from cognitive science – that of a mental model.

A mental model is an internal cognitive representation of some aspect of the real world (Johnson-Laird, 1983; Norman, 1983). As we navigate through life, our minds distill all of our encounters and experiences into an ever-growing and ever-evolving series of mental models, and we rely upon these models to guide our reasoning, decision-making, and behavior, and to help us form expectations about the world around us. Since the human brain constructs and refines mental models automatically and often without conscious effort, our minds are filled with an almost unimaginable number of these models (Jones et al., 2011). If you know, for example, how to type without looking at your keyboard, or how to use your television’s

remote control without looking at the buttons, it is because your mind has constructed mental models of the layout of those devices. If you can find the light switch in your bedroom in the dark, or can make your way to your favorite restaurant without consulting a map, it is because your mind has built spatial mental models to assist you. What's more, many of our mental models are not exclusively our own, but instead largely overlap with those of other people. Consider, for example, what would happen if people did not largely share a mental model of what constitutes acceptable behavior in social situations, or the chaos which would ensue if drivers did not possess a shared mental model of the rules of the road.

Even if we are not consciously aware of it, cognitive theory suggests that each of us who has used the web has also developed a mental model of what we consider to be agreeable web interface design. We simply cannot help it – our minds construct these models automatically. But what is the nature of these mental models? Further, while variations from person to person certainly exist, to what extent do people possess a general, shared mental model of web interface design? It is these questions that we examine in this investigation. Although answering these questions can help us to better understand user perceptions and expectations about web interface design, doing so is also a critical first step in a broader program of research aimed at revealing deep insights into why certain interfaces feel natural or intuitive to use, while many others seem poorly designed or difficult to use.

IDENTIFYING AND STUDYING MENTAL MODELS

While every web user has developed a mental model of normative web interface design, we still face two sizeable challenges if we want to describe and analyze those models. First, given that people find it difficult to describe their mental models (Rogers et al., 1992), how can we peek inside their minds and extract the information that we need? This is a challenging question, and indeed the paucity of effective tools and techniques for eliciting and describing mental models has been cited as a major impediment to research in this area (Jones et al., 2011). Second, how can we combine the mental models from many different individuals to produce a general scientific conceptualization of web interface design that we can evaluate and explore? To begin to address these challenges, we first need to consider the spatial nature of web interfaces.

A Geometric Perspective

Since every website is currently rendered on a flat, two-dimensional surface (such as a display screen), Euclidean geometry provides a natural framework for studying web interface layouts. Adopting a Euclidian approach provides access to myriad geometric concepts – including coordinate axes, spatial locations, distance, and direction – that can be usefully leveraged to describe the arrangement of the elements comprising a web interface. Further, people intuitively seem to understand the geometric nature of web interface design, and unconsciously construct mental models of web interfaces that are rooted in geometry. For evidence of this, simply consider your answer to the following question: *Where should the navigation menu appear on a web page?* If you are like most people in the western world, you probably believe that navigation menus should appear toward the top of the page on the left side, or perhaps a bit to the left of center. Now think more deeply about this reply – terms such as “top”, “left side”, and “left of center” are all points of reference in a two-dimensional geometric space. When considering the question above, your mind immediately consulted a two-dimensional visualization of a web page – your mental model of a web interface – and identified where the navigation menu was most likely to appear. After recognizing that people's mental models are strongly geometric in nature, we developed what we call the “blank screen method” for studying mental models of website design.

The Blank Screen Method

The blank screen method is an approach that researchers can use to overcome the challenges associated with eliciting and describing mental models that contain a strong geometric component, such as those associated with web interface design. The method involves presenting research subjects with an image of a blank computer screen, and then asking them to click on the screen at the point where they would expect a given interface element to appear. The process is repeated for a series of interface elements, with the geometric coordinates of the selected location being recorded each time. At the conclusion of this process, the researchers will possess a set of (x, y) coordinate points, each of which represents the subject's best judgment of where a specific interface element *should* appear. Together, the collection of point coordinates gathered from a research subject is a representation of the subject's overall mental model of how an interface should be designed. Repeating the process for many subjects yields a collection of (x, y) coordinates for each interface element, and these points can be combined to produce a dataset which can then be explored and studied using graphical and statistical techniques.

For the current study, we designed a web-based software system to perform the tasks described above. After collecting demographic data (age, gender, and level of web experience), the system asked subjects to identify the most likely screen locations for nine distinct web interface elements, namely: an advertisement, a copyright notice, a corporate logo, a login/logout feature, a navigation menu, a search box, the text content of the page, the page title, and the page's video content. The order in which these interface elements were presented was randomized for each subject to mitigate the

possibility of any ordering effects (Saris and Gallhofer, 2007). Our system also asked subjects to verify their selected screen location before proceeding to the next interface element in order to ensure that the data gathered were of the highest quality possible. Whereas past research has considered user web layout expectations for shopping, corporate, and news websites (Roth et al., 2010), and whereas our inquiry was focused on the possible existence of a general, shared mental model of web interface design, we intentionally avoided asking subjects to consider any particular type of website during the experiment. A screenshot of our software system employing the blank screen method is shown below.

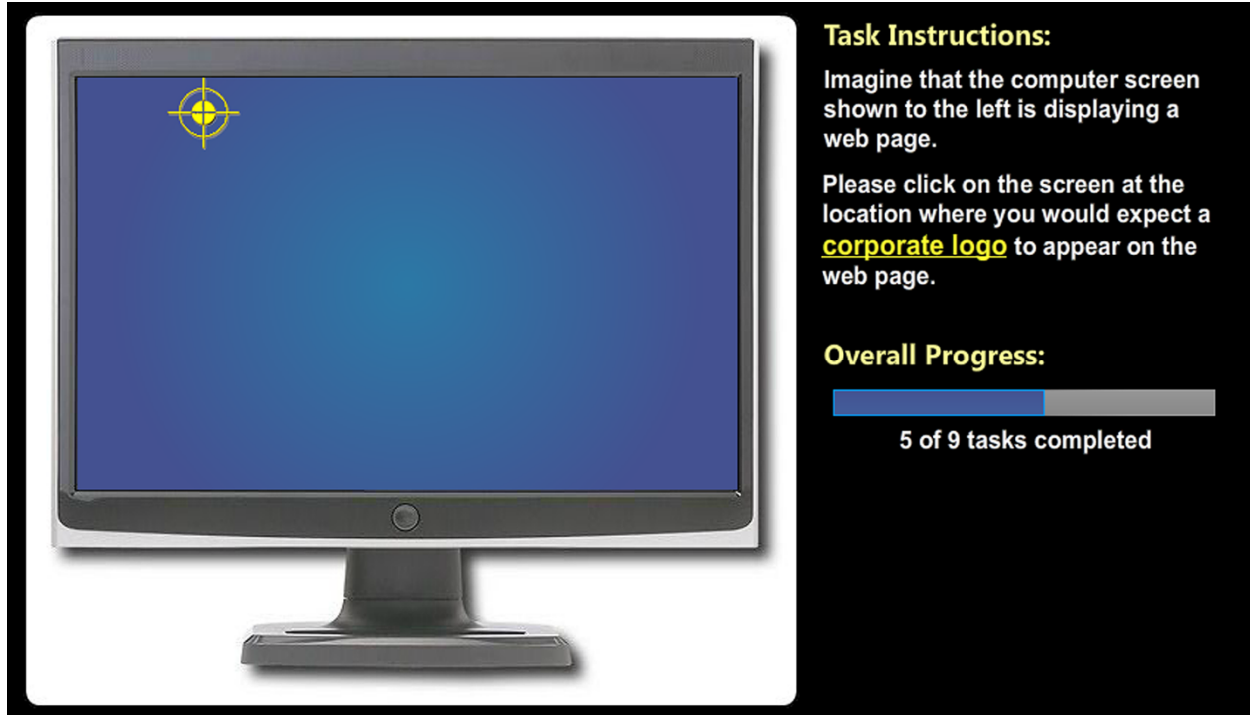


Figure 1. Screenshot illustrating the Blank Screen Method.

RESEARCH SUBJECTS, DATA, AND ANALYSIS

Since the target population for our investigation was adult web users in the United States, we engaged a major online marketing firm to craft a targeted advertising campaign for the purpose of soliciting volunteers for our research study. The firm's advanced technology allowed our advertisements to be displayed only to those subjects in the target population. IP address restrictions were also enforced to help prevent the same person from participating in the study more than once. In total, our campaign attracted 533 subjects, of whom 47.8% were male and 52.2% were female. Subjects reported an average level of web experience of 3.84 on a five-point Likert-type scale anchored at 1 = *very little experience* and 5 = *extensive experience*. Subjects ranged in age from 18 to 78 years. The average age was 32.9 years with an interquartile range of 25 to 37 years, indicating an age distribution skewed in the direction of youth. These values were consistent with overall Internet usage demographics among adults in the United States (Pew Research Center., 2012).

As noted previously, the primary data for the study were a series of (x, y) coordinates, each of which represented the screen location at which a subject indicated they would expect a given web interface element to appear. With 9 distinct interface elements and 533 subjects, we collected a total of 4,797 coordinate points. Analysis of these coordinate points and the underlying mental models that they represent was carried out using both graphical and statistical techniques. Since human beings are highly visual creatures, we began by transforming the coordinate points into a thermal image or "heat map" of each web interface element included in the study. These heat maps embrace the same principle as infrared thermal imaging, but instead of using different colors to show variations in temperature, they instead show the relative density of user clicks on different areas of the display screen, with "hotter" areas indicating greater density. The images which resulted from this process serve as graphical representations of people's mental models of web interface design, and are both fascinating and deeply insightful. Before discussing these heat maps, however, it remains for us to describe the statistical methods which were used in conjunction with the visual analysis.

The statistical methods used for analyzing the mental models data can be broadly divided into two categories: (1) clustering methods, and (2) exploratory methods. Hierarchical clustering was used to identify important clusters within the data for each of the web interface elements in the study, as well as to evaluate the degree of cohesion therein. This approach provided insight into the commonest locations selected by subjects for the various interface elements, and allowed us to quantify the optimal number of clusters per element. With this information, we were able to perform both inter-element comparisons and determine the extent to which a general, overall mental model exists among web users. The objective of the exploratory methods was to determine whether statistically significant differences were present in the locations of the various web elements according to a subject's age, gender, or level of web experience. For this purpose we employed multivariate analyses of variance (MANOVA), using the (x, y) coordinates of each interface element as dependent variables. Together, these statistical analyses provided a deeper understanding of people's mental models of web interface design than could otherwise be gained through visualization alone.

FINDINGS AND DISCUSSION

It has often been remarked that a picture is worth a thousand words, and in that spirit we shall begin the presentation of our findings with Figure 2 below, which contains heat maps depicting people's mental models of the locations of the interface elements investigated in the study.

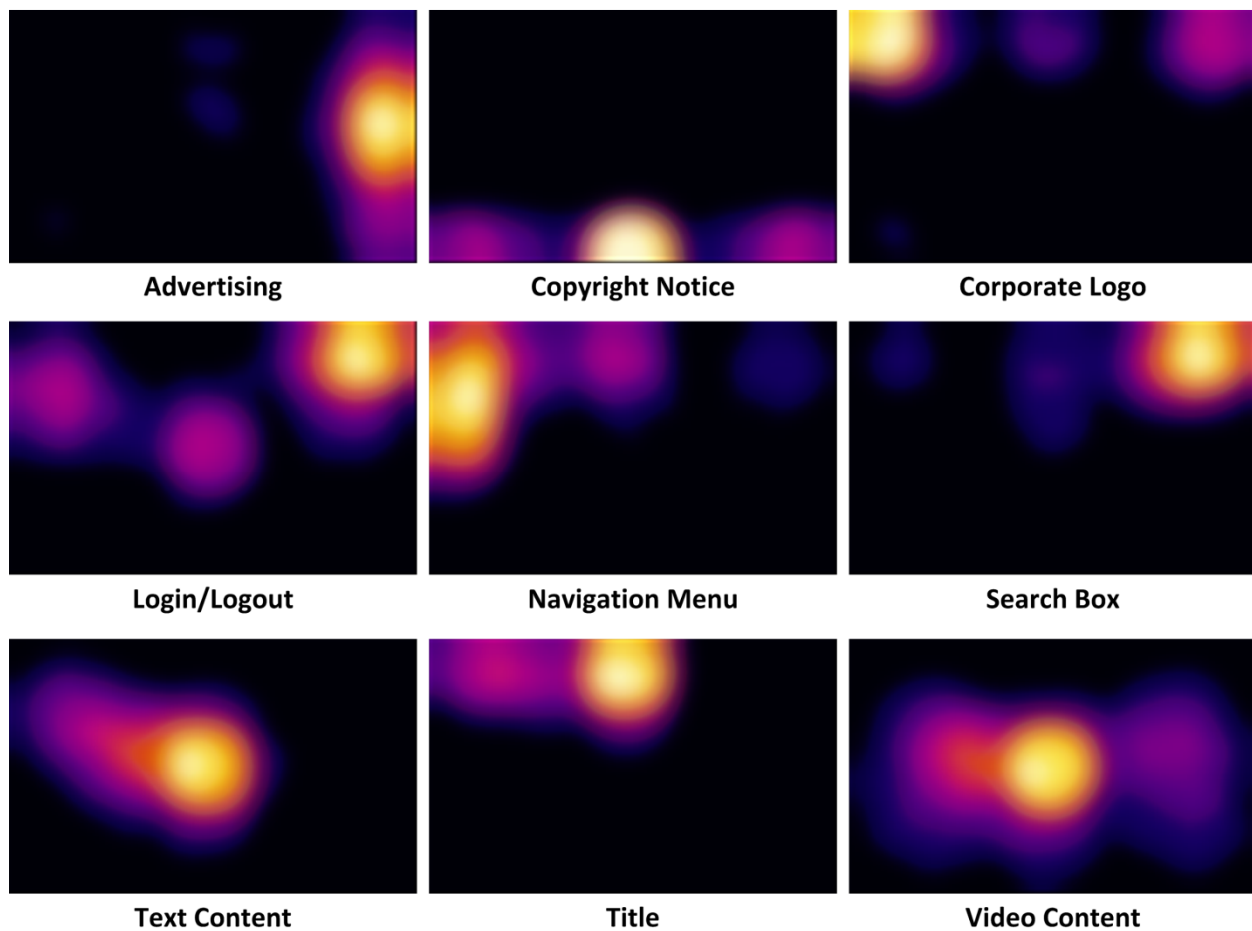


Figure 2. Heat maps depicting mental models of website design.

Each of the heat maps above depicts the relative density of clicks obtained from the research subjects for its corresponding interface element. In examining these heat maps, we were immediately struck by two phenomena. First, recalling that the heat maps were generated from the responses of more than 500 web users, there appears to be a remarkable degree of similarity among people's mental models of website layout. Put another way, there does not appear to be a great deal of variation from one person to another with respect to expectations about where interface elements should appear. On the contrary, web users seem to expect to find each interface element in one of only a very small number of locations on a web

page. Second, the heat maps show that each interface element appears to have one cluster which is dominant over any others. In light of the wide diversity in age, gender, and web experience of the subjects who participated in the study, this degree of cohesion seems quite extraordinary, and bodes well for the possible existence of a general, shared mental model among web users.

Cluster Analysis

Although the heat maps shown in Figure 2 appear to indicate that people's mental models of web interface design are reasonably consistent and cohesive from person to person, it was necessary to perform a cluster analysis on the data in order to quantify the apparent consistency and cohesiveness. Table 1 thus reports the number of major clusters for each web interface element, as obtained using average linkage-based hierarchical clustering. The table also provides the number and percentage of observations in each cluster, as well as the centroid coordinates of each cluster (as proportions of the width and height of the display screen). It is important to note that the results of the cluster analyses are not perfectly consistent with the distribution of user clicks depicted in the heat maps. These inconsistencies are attributable to the differences between the way in which human beings visually perceive the data and the quantitative view of the data, which is influenced by the constraints inherent in the clustering algorithm. The "Login/Logout" interface element serves as an effective example of this phenomenon. Whereas the heat map appears to show three clusters for this element, the clustering algorithm combines the two least populous of these three clusters into a single cluster, thus yielding a two-cluster statistical solution.

Interface Element	Clusters	Cluster	Observations (%)	Centroid (x, y)
Advertising	3	1	413 (78%)	(0.912, 0.502)
		2	76 (14%)	(0.491, 0.440)
		3	44 (8%)	(0.093, 0.600)
Copyright Notice	3	1	392 (74%)	(0.395, 0.955)
		2	128 (24%)	(0.844, 0.824)
		3	13 (2%)	(0.199, 0.309)
Corporate Logo	2	1	330 (62%)	(0.112, 0.199)
		2	203 (38%)	(0.775, 0.229)
Login/Logout	2	1	279 (52%)	(0.855, 0.163)
		2	254 (48%)	(0.315, 0.372)
Navigation Menu	3	1	475 (89%)	(0.186, 0.258)
		2	50 (9%)	(0.872, 0.268)
		3	8 (2%)	(0.166, 0.940)
Search Box	2	1	524 (98%)	(0.728, 0.163)
		2	9 (2%)	(0.081, 0.852)
Text Content	2	1	512 (96%)	(0.375, 0.454)
		2	21 (4%)	(0.790, 0.403)
Title	2	1	517 (97%)	(0.369, 0.135)
		2	16 (3%)	(0.791, 0.337)
Video Content	3	1	256 (48%)	(0.518, 0.285)
		2	165 (31%)	(0.236, 0.483)
		3	112 (21%)	(0.819, 0.526)

Table 1. Cluster analysis of web interface mental models data.

As suggested by the heat maps, the results of the cluster analysis confirm that each web interface element contains a very small number of clusters, and with the exception of the login/logout feature, each interface element has one cluster which clearly dominates the others. When considered together, these results point toward the presence of a general, shared mental model of web interface design. For purposes of the current study, this general model was characterized using the centroids of the dominant clusters for each interface element. By defining the shared mental model in this way, it became possible to compute the distance between each subject's personal mental model and the proposed shared mental model. Given that mental models are largely constructed on the basis of experience (Gentner and Stevens, 1983), and given that each person's experiences with the web are, in the aggregate, unique, each individual's mental model should reasonably be expected to diverge somewhat from a standard reference model. The relevant issue, then, is the extent of such divergence. We therefore calculated the distance between the locations selected by subjects for each web interface element and the centroid of that element's corresponding dominant cluster. After performing all of these calculations, we determined that in terms of the size of the display screen, each individual's mental model of web interface design varied from the reference model by an average of only 1.85%. This degree of proximity, we believe, provides powerful evidence for the presence of a general, shared mental model of interface design among web users.

Exploratory Analysis

To gain additional insights into the data, we next undertook an exploratory analysis using MANOVA. Initial omnibus tests (Rao, 1951) revealed that significant differences were present according to a web user's gender ($\lambda = 0.942$, $F = 1.740$, $p < 0.05$), age ($\lambda = 0.930$, $F = 2.126$, $p < 0.01$), and level of experience ($\lambda = 0.945$, $F = 1.661$, $p < 0.05$), thus justifying a more granular investigation. Further analyses revealed that only the location of video content varied significantly by gender, indicating that men and women have slightly different expectations regarding where video content is expected to appear on a web page. Specifically, in terms of the width and height of the display screen, women tend to expect video content to appear approximately 3.5% to the left and 4.3% lower on the screen than do men. With respect to age, significant differences were detected in the locations of a web page's text content and navigation menu. These differences were revealed to be almost entirely attributable to the expectations of the older subjects in the study. Those subjects who were at least 60 years old expected a web page's text content to appear approximately 7.4% of the screen height closer to the top than younger web users. These older subjects also expected the navigation menu to appear approximately 9.2% of the screen width farther to the right and 10.5% of the screen height closer to the top when compared to younger web users. Finally, the expected locations of an advertisement, a login/logout feature, and the navigation menu varied significantly according to a user's level of web experience. These differences were due almost entirely to the expectations of the least experienced web users. Again in terms of the width and height of the display screen, the least experienced users expected advertisements to appear approximately 7.9% farther to the left and 3.9% higher than a user of average experience. Notably, the least experienced users also expected the login/logout feature to appear 54.1% farther to the left and 19.9% higher than a typical web user, while expecting the navigation menu to appear 65.9% farther to the right and 19.3% higher on the screen. These observations accord with cognitive theory, which suggests that mental models are unstable and evolve over time as one accumulates more experience (Johnson-Laird, 1995).

LIMITATIONS AND CONCLUDING REMARKS

There are several limitations to our work which merit acknowledgement. First, our findings were derived only from the responses of adult web users. Given the role of experience in the formation of mental models, we should expect children's mental models to differ from those of adults. Indeed, investigating children's mental models may reveal much in regard to both the design of intuitive interfaces and the evolution of mental models of interface design over time. Second, our results are similarly limited by the fact that our subjects were all English-speaking web users. This is important because the English language is written from left-to-right, and mental models of the locations of peripheral interface elements such as a navigation menu, advertisement, or corporate logo might reasonably be expected to differ among web users who speak Hebrew, Arabic, or Farsi, which are written from right-to-left. These limitations represent fertile ground for future research.

The limitations noted above notwithstanding, there are three key contributions of this work which we would like to emphasize. Recalling that our first research question inquired into the nature of mental models of web interface design, the results of our study suggest that people are readily able to describe their internal cognitive representations of web interfaces in geometric terms, and we therefore conclude that mental models of web interface design are strongly geometric in nature. From a cognitive perspective, this implies that web users unconsciously componentize web interfaces into a collection of constituent elements (e.g., an advertisement, a title, a search box, etc.), and then establish spatial expectations about the locations of those elements which are rooted in geometric relativism. Next, our second research question addressed the possible existence of a general, shared mental model of web interface design. Our results reveal not only that people possess a shared mental model of web interface design, but also that this shared model is remarkably cohesive and consistent from

person to person. Given the human brain's proclivity for generating mental models, it stands to reason that web users develop a wide array of sub-models, including specific models for the websites that they visit most frequently, and models for different types or categories of websites (e.g., shopping sites, news sites, etc.). Indeed the presence of such sub-models has been empirically demonstrated by past research (Roth et al., 2010). From the perspective of cognitive theory, these sub-models can be viewed as descendants of the general mental model discussed herein. Although such sub-models possess distinguishing characteristics, their features are largely derived or inherited from a more general mental model. This general model reflects the collective expectations of web users, and may serve as a useful point of embarkation from which to investigate how the degree of alignment between a mental model and a given web interface affects user perceptions about the design and usability of that interface.

The final contribution of this work is the blank screen method, which, although not tied directly to our research questions, nevertheless deserves a few additional remarks. Our study shows that the blank screen method can serve as a valuable and effective research tool for eliciting and describing mental models that contain a strong geometric component. Although the method was developed here for use in the context of web interface design, it is by no means limited to such a narrow domain. In the spirit of crowdsourcing, and inasmuch as the collective wisdom of a group can often outperform that of an individual expert (Surowiecki, 2005), the blank screen method could be readily adapted to allow large groups of people to collaboratively design the layouts of a wide variety of spaces, including parks, stores, museums, cities, etc. In light of the noted paucity of tools for eliciting and studying mental models (Jones et al., 2011), it is our sincere hope that researchers working in this area will find the blank screen method to be both an illuminating and valuable addition to their methodological toolkits.

REFERENCES

1. Agarwal, R. and Venkatesh, V. (2001) Assessing a Firm's Web Presence: A Heuristic Evaluation Procedure for the Measurement of Usability, *Information Systems Research*, 13, 2, 168-186.
2. Flavián, C., Guinalú, M., and Gurrea, R. (2006) The role played by perceived usability, satisfaction and consumer trust on website loyalty, *Information & Management*, 43, 1, 1-14.
3. Gentner, D. and Stevens, A. L. (1983) *Mental Models*. Hillsdale, NJ: Lawrence Erlbaum Associates.
4. Johnson-Laird, P. N. (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, UK: Cambridge University Press.
5. Johnson-Laird, P. N. (1995) Mental Models, Deductive Reasoning, and the Brain, in *The Cognitive Neurosciences*, Cambridge, MA: MIT Press, 999-1008.
6. Jones, N. A., Ross, H., Lynam, T., Perez, P. et al. (2011) Mental Models: An Interdisciplinary Synthesis of Theory and Methods, *Ecology and Society*, 16, 1, 46.
7. Nadkarni, S. and Gupta, R. (2007) A task-based model of perceived website complexity, *MIS Quarterly*, 31, 3, 501-524.
8. Norman, D. A. (1983) Some Observations on Mental Models, in D. Gentner and A. L. Stevens (Eds.) *Mental Models*, Hillsdale, NJ: Lawrence Erlbaum Associates, 7-14.
9. Palmer, J. W. (2002) Web Site Usability, Design, and Performance Metrics, *Information Systems Research*, 13, 2, 151-167.
10. Pew Research Center. (2012) *Who's Online: Internet User Demographics*. Washington, D.C.: Internet & American Life Project.
11. Rao, C. R. (1951) An asymptotic expansion of the distribution of Wilks' criterion, *Bulletin of the International Statistical Institute*, 33, 2, 177-180.
12. Rogers, Y., Rutherford, A., and Bibby, P. (1992) *Models In the Mind: Theory, Perspective, and Application*. London: Academic Press.
13. Roth, S. P., Schmutz, P., Pauwels, S. L., Bargas-Avila, J. A. et al. (2010) Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages?, *Interacting with Computers*, 22, 2, 140-152.
14. Saris, W. E. and Gallhofer, I. N. (2007) *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, NJ: John Wiley & Sons.
15. Surowiecki, J. (2005) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York, NY: Anchor Books.