Brand Awareness and the Evaluation of Search Results

Bernard J. Jansen	Mimi Zhang	Ying Zhang
College of Information Sciences & Technology	College of Information Sciences & Technology	College of Engineering
The Pennsylvania State University	The Pennsylvania State University	The Pennsylvania State University
University Park, PA 16802	University Park, PA 16802	University Park, PA 16802
jjansen@acm.org	mzhang@ist.psu.edu	yzz114@psu.edu

ABSTRACT

We investigate the effect of search engine brand (i.e., the identifying name or logo that distinguishes a product from its competitors) on evaluation of system performance. This research is motivated by the large amount of search traffic directed to a handful of Web search engines, even though most are of equal technical quality with similar interfaces. We conducted a laboratory study with 32 participants to measure the effect of four search engine brands while controlling for the quality of search engine results. There was a 25% difference between the most highly rated search engine and the lowest using average relevance ratings, even though search engine results were identical in both content and presentation. Qualitative analysis suggests branding affects user views of popularity, trust and specialization. We discuss implications for search engine marketing and the design of search engine quality studies.

Categories and Subject Descriptors
H.3.3 [1] Information Search and Retrieval – *Search process*: Measurement, Experimentation, Human Factors

General Terms: Performance, Design, Economics, Human Factors

Keywords: Brand, Web searching, search engines

1. INTRODUCTION

Since its inception, there has been a rapid growth in the search engine market. Web search engines continue to attract large number of Web searchers, and they consistently rank as some of the heavily visited sites in the market in terms of the number of visitors. With thousands search engines on the Web, only a handful search engines dominate in terms of usage. Why?

Viewed from a technological perspective, this clustering is interesting because studies report that the performance of most major search engines is practically the same [c.f., 1]. Performance is defined as returning relevance results. Performance is measured by precision, which is the ratio of relevant documents to the total number of documents returned at some point in the results listing. The interfaces of most search engines are also similar, namely a text box, some verticals (i.e., tabs for searching the Web, Images, Audio, etc.), and some links to view result pages.

With the similarity in terms of technology and interface design, why do only a small number of search engines dominant Web traffic? Do other elements affect the evaluation of a search engine's performance? Seeking the answers to these questions motivate our research.

2. REVIEW OF LITERATURE

The interfaces of Web searching engines contain branding elements, such as logos, and names. A brand is the intangible sum of an organization's attributes, which can include an organization's name, history, reputation, and advertisement. A brand is also an identifying symbol, sign, or name that distinguishes an organization or a product from its competitors. Good branding can results in loyal customers. However, the effect of branding on technology design is not well studied.

Park, Harada, and Igarashi [2] report that the users' perceptions of a product's brand affect their perceptions of mental demand. While there may be some recognition that branding is important in the marketing of product, there has been little research into the brand effect on the evaluation of system performance. In this research, we measure the effect of brand perception on user perception of the performance of Web search engines.

3. RESEARCH OJECTIVES

Our research objective is: *How does branding affect overall user evaluation of results retrieved by Web searching systems?* In order to address this research question, we designed a study that altered the brand of search engines for a set of queries while controlling for the quality and display of the results.

4. RESEARCH DESIGN
4.1 Data Preparation

We first extracted a set of e-commerce queries from an approximately one and half million queries Web search engine transaction log using a modified snowball technique. From these queries, we selected four queries representing four searching domains: medical, entertainment, travel, and ecommerce. We developed searching scenarios around each of the four queries. The four queries used were: *camping mexico*, *laser removal*, *manufactured home,* and *techo music*.

We then submitted these four queries to Google, a major U.S. search engine using a software application that submitted the queries and retrieved the first search engine results page (SERP) for each query exactly as it would be presented to a human user. The total time from submission to completion of result retrieval took approximately 30 seconds. We then removed all identifying logos, text, URLs, and HTML code from the Google result pages. We removed the redirects in the results, so the URLs pointed directly to the targeted Web site. This left us with four cleaned results pages.

Table 1. Comparison of Average Precision Scores by Query and by Search Engine

Search Engines	Queries				Average	Difference (Average
	camping mexico	laser removal	manufactured home	techo music		
AI^{2}RS	0.35	0.31	0.26	0.37	0.32	-10.3%
Google	0.26	0.25	0.69	0.27	0.36	0.7%
MSN	0.44	0.29	0.30	0.34	0.34	-5.7%
Yahoo	0.39	0.29	0.55	0.44	0.42	15.3%
Average	0.36	0.28	0.45	0.35	0.36	0.0%

We then got screen captures of SERPs from Google, MSN Live Search, and Yahoo!, all major and well-known Web search engines, for each of the four queries. We developed an in-house search engine, AI^{2}RS, and got screen captures of the AI^{2}RS results pages for each of the queries. Using the cleaned Google results and the images from the AI^{2}RS, Google, MSN Live Search, and Yahoo!, we developed four experimental SERPs for each of the four queries. At the end of this process, we had sixteen experimental SERPs, four from each search engine for each of the four queries. Regardless of the search engine branding elements, the results were identical across all search engines for each query.

4.2 Study Procedure

We recruited 32 participants from a major US university. The age range was 18 to 25 years. We presented each participant with all four queries scenarios, one at a time. Each participant completed one query before moving to the next. The moderator would read the applicable scenario before moving to the next query. We counterbalanced the order of search engines and the order of the searching scenarios to eliminate ordering effects. While the participant was searching, the moderator annotated utterances and user actions using an application designed for quantitative and qualitative data capture during Web searching studies. After the participant had completed all four query sessions, the moderator returned the participant to the first query, and the participant visited all Web pages for each query that the participant had not visited during the session. The participant evaluated the Web document and presented a basis for the evaluation. Approximately one hour was required to complete the sequence for each participant.

5. RESULTS

Returning to our research question (*How does branding affect overall user evaluation of results retrieved by Web searching systems?),* results are shown in Tables 1. Of all the search engines, Yahoo! had the overall best average precision of 0.42. This was 15.3% better than the average of all four search engines. The overall average for all search engines over all queries was approximately 0.36. The spread among search engines was 25%. The implications of these research findings give empirical weight to the notion that affective and cognitive user perceptions affect user interaction with systems and interactions. Therefore, product brand is an important usability variable in system design and evaluations.

Apart from our quantitative results, our qualitative analysis show that users are sensitive about branding, and branding has various meanings. Branding denotes *favorite*, *familiarity* and *popularity* to users. When participants commented on their frequently used search engines, they said "*I am familiar with Google.*"; "*Google*

is a pop culture terms; this is why Google is so popular."; "*I go to Yahoo! a lot for searches in English*". Branding also indicates **trust**. As an example, when searching for medical information, the participant tried to explain why she checked some results rather than the others. She uttered "*Since this is about health issue, I will look for good Websites. I would go to a doctor or company that I have heard of and trust.*" The underlying intent of her words is that a doctor or company Website she has heard of before would have a good, trustworthy, and positive branding image. Branding also means **specialization**. When confronted with certain tasks, the users will immediately think about these sites. When seeing the branding elements of these sites, they make instant assumptions about what they are like. Websites like Amazon and eBay mean ecommerce. Searchers believe they need pay to get useful information. When talking about music, the users would think about iTune or Napster right away. On the other hand, the negative branding image has unfavorable effects and leads to the loss of trust from users. A participant told us that he never goes to AOL because "*AOL is inferior to the others in my mind.*"

6. DISCUSSION

Study findings show that branding as a perception of product has a dramatic effect on user's evaluation of system results. Performance evaluations varied by more than 25% between the top-most rated search engine and the bottom even though results were identical in both content and in presentation. It appears that even though Google is the most commonly used engine for searching, Yahoo! has a positive branding awareness. This may help explain why Yahoo! has endured and prospered in a competitive marketplace where so many other search engines (i.e., Excite, Northern Light, and Infoseek) have come and gone. Future research involves quantitative and qualitative analysis of experimental data to tease apart the nuanced relationship between perception of performance and brand, and how to incorporate branding into the system design process.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Eastman, C. M. and B. J. Jansen, "Coverage, Ranking, and Relevance: A Study of the Impact of Query Operators on Search Engine Results," *ACM Transactions on Information Systems*, vol. 21, pp. 383 - 411, 2003.

[2] Park, S., A. Harada, and H. Igarashi, "Influences of personal preference on product usability," in *Proceedings of Conference on Human Factors in Computing Systems: CHI '06 extended abstracts on Human factors in computing systems,* pp. 87 - 92.