An Empirical Analysis of Internet Search Engine Choice

Rahul Telang (rtelang@andrew.cmu.edu)

Tridas Mukhopadhyay (tridas@andrew.cmu.edu)

Ronald T. Wilcox (wilcoxr@darden.virginia.edu)

Send correspondence to

Rahul Telang

H. John Heinz III School of Public Policy and Management

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, PA 15213-3890

rtelang@andrew.cmu.edu

December 2001

An Empirical Analysis of Internet Search Engine Choice

Abstract

We investigate consumers' choice behavior for Internet search engines. Within this broad agenda, we focus on two interrelated issues. First, we examine whether consumers develop loyalty to a particular search engine. If loyalty does indeed develop, we seek to understand the role of loyalty in the search engine choice. We also explore how the use of non-search features such as email, news, etc. provided by the engines enhances or inhibits customer loyalty. Second, we seek to determine how search engine performance affects the user choice behavior. To accomplish our research objective, we first develop a conceptual model of search engine choice based on the literature of human-computer interaction and cognitive psychology. Our model reflects the fact that information goods such as search engines are a fundamentally different class of products than common household items. We posit that the ability to learn various search engine features and the ease (or difficulty) of transferring this learning to other engines would determine loyalty in this context. Indeed, we expect the user to exhibit differing levels of loyalty to search and non-search features of the engines. We also expect that dissatisfaction with search results would negatively affect search engine choice. Next, we use a multinomial logit model to study 6,321 distinct search engine choices for six engines over a period of one year. Our findings show that user dissatisfaction with search results has a negative effect on both immediate and future user choice decisions. We also find that the impact of loyalty is small when consumers use engines primarily for search purposes, but it is quite large when they use personalized features. The results of this research provide insight into consumer behavior in the marketplace for Internet search engines, and offer guidance to managers in product development.

1. Introduction

Internet search engines often act as the gatekeepers of electronic commerce. Consumers may begin their sessions on the Internet by visiting a search engine, and move on to gather information, browse products, or make purchases only after the Web sites of interest have been located through the search procedure (Kraut et al., 1996). The central role of search engines in disseminating Web site information has left the companies that provide search engines in a powerful position in this marketplace. According to Media Metrix (October 2001), they are some of the most visited sites on the Internet and the most attractive vehicle for online advertising (Target Marketing 2000).

Companies like Yahoo!, Excite, Lycos, and other portal sites expect that they will be able to achieve a sustainable competitive advantage by acquiring and cultivating long-term repeat users (*Fortune*, 1999). By generating loyalty for their search engines, these companies can capture revenues by pairing their search engine, often through company acquisitions, with other content-oriented commercial sites (*Money,* 1999). A loyal customer is worth more to Yahoo! than a non-loyal one because repeated interactions with the portal and potential product purchases on sites linked to the portal, provide Yahoo! with a rich set of information about preferences and purchase patterns at the individual level. Individual-level data can then be used by Yahoo! directly or sold to other companies to tailor product offerings and even pricing plans to identifiable and targetable consumers (Petrison et al. 1997). Although the business model relies on the fact that consumers will be repeatedly choosing their favorite search engines, it is not clear as to what might lead to such repeat use or how consumers interact with the engines.

The purpose of this research is twofold. First, we wish to model consumer choices for Internet search engines, and examine how important loyalty is in search engine choice. We also explore how the use of non-search features such as email, news, etc. provided by the engines enhances or inhibits customer loyalty. Second, we seek to determine how search engine performance affects the user choice behavior. The results of this research will provide insight into consumer behavior in the marketplace for Internet portal and similar information goods, offering guidance to managers of these companies for making better product design decisions.

We develop a conceptual model of search engine choice based on the literature of human-computer interaction and cognitive psychology. Our model reflects the fact that information goods such as search engines are a fundamentally different class of products than common household items. Search engines are freely available, and easily accessible on the Internet. Why would then a user repeatedly return to the same engine? We posit that the ability to learn various search engine features and the ease (or difficulty) of transferring this learning to other engines would determine loyalty in this context. Indeed, we expect the user to exhibit differing levels of loyalty to search and non-search features of the engines. We also expect that dissatisfaction with search results would negatively affect search engine choice.

Our analysis is based on Internet navigation data from 102 demographically diverse users over a period of one year. We use a multinomial logit model to study 6,321 distinct search engine choices. Since the users in our sample differ clearly in their switching propensity, we attempt to account for this heterogeneity. For example, some users never exhibit intra-session switching behavior. We use the Mover-Stayer (MS) structure in our model to account for this type of heterogeneity. Our results show that while loyalty is indeed a significant factor in a user's choice decision, it also depends on the use of non-search features. Our results show strong repeat use for personalized features but the same is not true for the search function alone. In addition, we find that the ability of the engine to satisfy the user with its results plays a strong role in engine choices. Our findings show that user dissatisfaction with search results has a negative effect on both immediate as well as future user choice decisions.

The rest of this paper is organized as follows. In the next section, we discuss the prior research on loyalty. In Section 3, we present the conceptual model underlying our analysis, and discuss the formal hypotheses to be tested. Section 4 links the choice model described in Section 3 to data on search engine choice behavior. Section 5 details the results of our estimation.

Finally, Section 6 concludes our paper with some managerial implications of our research and points to directions for future research.

2. Prior Research

Often, loyalty models are employed to study the choice behavior among many alternatives. In the case of consumer choices for jobs, household goods, or transport, the effects of various explanatory variables have been studied extensively (Ben-Akiva and Lerman 1985). It has also been clearly demonstrated, especially in marketing literature, that brand loyalty along with other variables often acts as an important determinant of product choice (e.g., Guadagni and Little 1983). Recently, Smith and Brynjolfsson (1999) and Chen and Hitt (2000) modeled user choices for online book and brokerage services respectively. However, the same cannot be said about the market for Internet portal sites. Indeed, within the broad context of Internet portal sites, we know of no studies that have modeled consumer choices. We believe that portal sites and, in particular, search engines are in several important respects a fundamentally different class of product than what consumers experience in their day-to-day shopping and consumption behavior.

Prior research in consumer choice models has used various explanatory variables along with loyalty for explaining the particular choice made. Typical economic variables studied include price, coupons, etc. Given that search engines are free, there are hardly any economic reasons driving their choices. Unlike any other product studied before, search engine use involves a lot of interactions with them. We believe that consumer choices will actually depend on *how* consumers interact with the engines rather than any other external factor. Depending upon the way users interact with the engine, and how satisfied or dissatisfied they are with the interactions, more than anything else, will determine the immediate as well as the future use of the engines, and hence their loyalty. Unlike other goods, where it is difficult to track switching behavior due to dissatisfaction with the product, interactive information goods like search engines allow us this opportunity. Thus we can test the effect of such interactions on user choices.

Prior research has also demonstrated the positive effect of past choices on current choices. A common explanation for brand loyalty is “inertia,” where consumers do not even consider other available brands (Jeuland 1979). Shugan (1980) conjectures that this type of routinization reduces the “cost of thinking,” and hence leads to repeat purchase. Another related explanation concerns consumer uncertainty about other brands. From this uncertainty flows the possibility of economic and psychological costs of switching brands (Hawkins et al. 1986). Economic costs arise from the possibility that the newly purchased brand is not of expected quality, and therefore not worth the price paid for the item. Potential psychological costs include the discomfort associated with using an item that does not perform up to expectations (e.g., eating a poor tasting cereal), as well as the possible social embarrassment arising from product failure (e.g., using a shampoo which does not work well or drinking a brand of beer which is not appropriate in a given social context).

Consumers develop a trusting relationship with their favorite brands that leads to repetitive purchase behavior. This trust mitigates the risk associated with product purchases and provides assurance that the product will indeed perform as expected. Erdem and Keane (1996), for example, propose that consumers learn about the brand of choice through usage experience and that usage experience leads to a better understanding of the brand and lower risk perceptions.

While some of the explanations for brand loyalty for consumer goods may apply to search engine choice, we believe that there are two strong reasons that will lead to the repeated use of the engine. The first reason concerns the ability of the engine to satisfy the user. The second reason relates to the learning costs incurred by the user and the ability of the user to transfer this learning to other engines. Users constantly interact with the engines. They not only learn the interface of the engine but also develop expectations about the quality of the search results. Moreover, search engines also offer many non-search features that affect the transfer of learning and hence the strength of loyalty.

Search engine companies attempt to lock in users by adding new features to their products. Yahoo!, for one, spends close to fifteen percent of its annual budget on product development (Annual Report, 2000). Offering value-added features in software is certainly not a new idea. Nault and Dexter (1995), for example, show how adding value-added features to software enables firms to charge a premium to their customers. There have been some attempts in the IS literature to measure the impact of these features. Brynjolfsson and Kemerer (1996) study the spreadsheet market while Rao and Lynch (1995) examine the case of computer workstations. Both studies employ the hedonic price analysis approach. Search engines typically include value-added features such as personalized news, email, and chat rooms that allow the user do much more than just search. The issue of personalization on the web and its benefits has attracted a lot of attention (Gilmore et al. 2000). The technology behind personalization has also been studied extensively (Mulvenna et al. 2000; Mobasher 2000). Some metrices to measure the impact of such programs have also been proposed (Schonberg et al. 2000). However, no rigorous empirical analysis with actual usage data has been undertaken to understand the role of personalization in loyalty formation on the Web. Studying the effects of search engine features is particularly interesting because we believe that different features impose different switching costs and affect the level of loyalty in different ways.

3. Conceptual Model

One might argue that users need not be loyal to any search engine because they can easily switch to a competing engine without incurring any costs. In popular terms, another engine is only a mouse click away. However, we argue that this is only partially true. A positive experience with an engine might lead to repeated use even though other engines are available. People may become loyal to specific search engines because of knowledge gained about the search engine through prior usage experience. Consumers must spend time in learning to use search facilities and a variety of value-added features offered by the search engines. While learning imposes switching costs, unsatisfactory search results and the ability to transfer learning to another engine encourage users to switch.

In this section, we lay out our conceptual model for search engine choice. We then develop formal hypotheses from this conceptual model, and test these hypotheses using data from a cross-section of Internet users.

3.1 Mental model of loyalty through learning

The literature in human-computer interaction and cognitive psychology has shown that there is considerable exploratory learning involved in using computer interfaces (Carroll 1987). Users often develop mental models about the system and then attempt to apply these models to increase task efficiency when new tasks arrive (Johnson-Laird 1983; Gentner and Stevens 1983; Staggers and Norcio 1993). These mental models develop out of various learning activities in which users engage. The specific model that a given user forms can have a significant impact on future task efficiency (Lim et al., 1997). For example, anecdotal evidence suggests that as people get used to a general "look and feel" on the Web, comprehension time reduces, and the ability to process information substantially increases (Austin 1997). Thus, the success of the learning process in increasing task efficiency can be an important determinant of subsequent system utilization (Davis and Bostrom 1993).

Repeated interactions with the engine can lead to a comfort level with it and therefore form loyalty. However, there are two reasons that can weaken this switching cost. First, learning can be transferable, thus making it easy for the user to switch. Second, the search results may not be satisfactory, leading the user to switch to other engines.

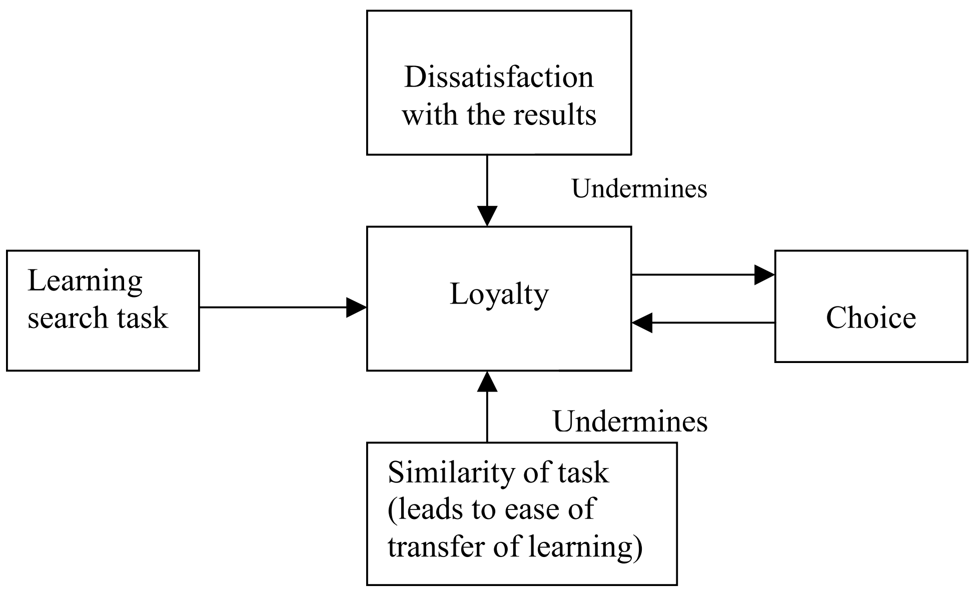
Learning in this context can occur in two stages. The first stage entails learning the general pattern of using the search function. We believe learning to use the search function is sufficiently generic, and can be transferred to any engine. The second stage involves learning the specific search engine features such as news and email. The switching cost to a great extent depends on how easily this learning can be transferred. The amount of this transfer actually depends on how similar the task is on two different engines (Carroll 1987, Newell et al. 1983, John and Kieras 1996).

The GOMS (Goals, Operators, Methods and Selection Rules) Model (Newell et al., 1983) has been widely applied to quantify the measure of learning and transfer of learning. According to this model, a task can be decomposed into perceptual, cognitive, and motor activities (Lim et al. 1996). Perceptual activities include reading information and locating an icon, while cognitive activities include remembering mapping instructions and icons, formulating goals, and developing an action plan. Motor activities include moving the mouse and clicking, for example. In most instances, these activities can be performed in parallel (John and Kieras 1996). Taken in totality, these observations of human behavior directly imply that as a user becomes more comfortable with an interface, she also becomes more efficient using it.

Based on this theory, tasks that are similar in nature would make learning easily transferable and would have low switching costs. Or, tasks that are different would make learning difficult to transfer, and would have high switching costs. For the purpose of a simple search, all engines are quite consistent in their designs. The search feature (which involves a text box and a search button) is easily visible, and easy to locate on the popular engines. All of them require the user to enter the search term(s) and click the "search" button. This task is quite similar across engines in terms of perceptual, cognitive, and motor activities.

However, due to the inherent properties of search engines, their results may not always satisfy the users (Bradlow et al., 1999; Mukhopadhyay, Rajan, and Telang 2001). Depending on the engine characteristics, search terms, and the user herself, an engine may sometime dissatisfy her with its results. This in turn may lead the user to switch to another engine immediately and start the search process anew. While for traditional goods it is difficult to detect whether a switch occurs due to dissatisfaction, server records of search engine use can detect this phenomenon with a high degree of confidence. Whenever a user switches to another engine but continues the process with a similar search term, it suggests that the user is not entirely satisfied with the results. This act of switching due to dissatisfaction is an erosion of loyalty formation. The user now learns the alternate engine through appropriate perceptual, cognitive and motor activities. Assuming the alternate engine provides satisfactory results, this experience breaks down switching barriers, and makes it likely that the engine chosen as a "back-up" choice for the current search task may be chosen first in subsequent search tasks. Repeated negative experiences with an engine exacerbate this effect. Figure 1 depicts the conceptual foundations of our model of loyalty.

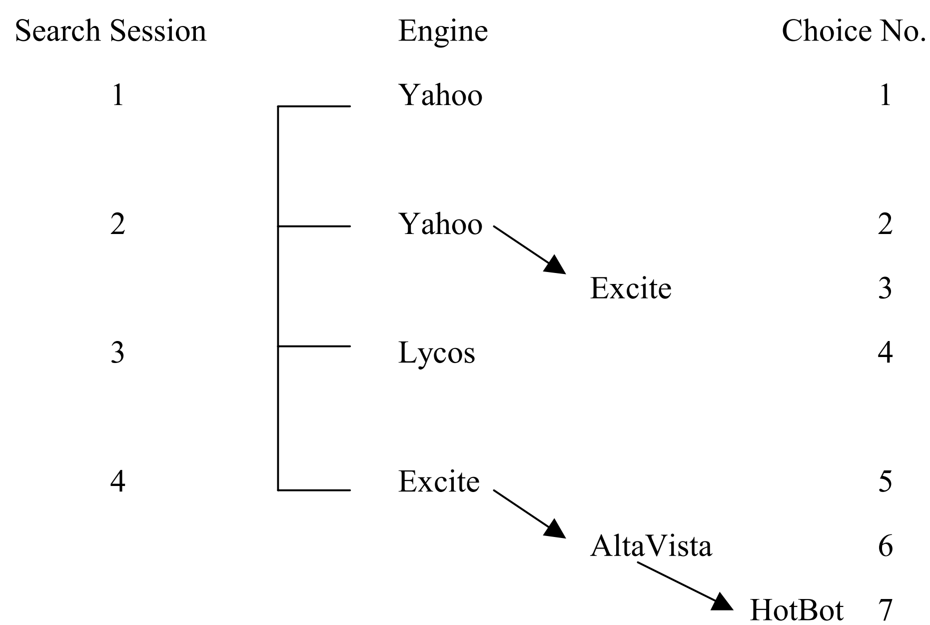
Figure 1. Conceptual Model of Search Engine Choice



Consider the following choice sequence for a user (Figure 2). The user has chosen seven engines in four search sessions. In the first search session, the user visits Yahoo!, performs a search and quits. Next time, she again starts her search session with Yahoo!, but switches to Excite to continue the same search. We conjecture that this immediate switching occurs due to dissatisfaction with the results. Here, Excite represents an "intra-session switching" as it occurs during the same search session. In this example, the user visits Lycos for her third search session. Finally, in the fourth search session, the user starts with Excite, becomes dissatisfied with the results, and switches to Altavista to continue the same search. She is dissatisfied again with the Altavista results and switches to HotBot to continue the search.

The choices of both Altavista and HotBot in Figure 2 represent intra-session switching. This switching to another engine is enabled by the ease afforded by the browser interface, and the free services offered by the engines. This type of switching to a different product for the same consumption occasion is not common for traditional goods. For example, a consumer would rarely drive to a restaurant, order a meal, and then discard it, only to drive to another restaurant. More commonly, she would complete the meal, but would try another restaurant the next time. This latter switching behavior occurring at a different consumption occasion is comparable to the switch from Lycos to Excite (choice number 4 and 5) in our example. As pointed out, there is a difference between intra-session switching (which is for the same search task) and switching later on for a different search task. We argue that since this intra-session switching from a particular engine is due to dissatisfaction with its results, it leads to poor perception of the engine.

Figure 2. Choice sequence for a user



The dissatisfaction from poor results affects the choice behavior in two ways. First, poor results from the engine may lead to intra-session switching. Second, it may weaken user loyalty for future choices. We define two measures to capture these effects. The first measure is defined as the occurrence of intra-session switching (OS). We conjecture that after every search engine choice, users may switch to another engine with a certain probability within the same search session. For any choice *k*, OS takes the value of one for the engine chosen at choice (*k*-1) and zero otherwise. By definition, the user would choose a different engine for choice k as she makes an intra-session switch with a certain probability. OS also accounts for the fact that intra-session switches are not necessarily independent choices.

The second measure captures the effect of dissatisfaction on future choices, and is termed cumulative negative experience (CE). As the name suggests, this is the proportion of times the user made an intra-session switch away from an engine to the total number of times she used the same engine. In Figure 2, the user chose Yahoo! twice (choice 1 and 2) and switched away once (choice 3). Therefore for a future choice (e.g., choice 4), CE will be equal to half for Yahoo!. It is clear that CE for Yahoo! may change over time.

Based on the forgoing discussion, a loyal consumer base will emerge only if the engines offer high-quality results consistently. Otherwise, lower switching costs and unsatisfactory results will force switching and erode loyalty.

3.2 Non-Search Features

Our focus, until now, has been on the search function. But users visit search engines also for many non-search features. Users check stocks, use email, participate in chat, and so on. While switching is relatively easy for search tasks, the use of non-search features may raise barriers to switching. We first define the non-search features, and then explain why they may engender loyalty in this market. We divide them into two categories.

Non-personalized features: As the name suggests, these features do not require personalization. We define a non-personalized feature as any service the user can access directly from the search engine, and does not have to enter a username and password. There are many examples of this type of features such as news, weather, sports, etc. For example, users can often access current news by clicking the appropriate link on the search engine site. Using this service requires no registration by the user and no entry of her name or password.

Personalized features: These features require registration, and entry of a username and password to access the service. At the same time, they let users customize their interactions with the site. All the search engines we examined offer personalized services for emails, chat rooms, community bulletin boards, messaging services, and personalized home pages. In addition to meeting the aforementioned conditions, all of these features have interfaces of their own. For example, when a user accesses Yahoo! mail, she gets a completely different interface from the original search engine interface.

Learning non-personalized features is cognitively simple as it merely involves looking up the feature and clicking it. However, perceptually it is difficult to transfer this learning because it involves knowing and identifying the feature location on the engine interface. Size, location, color, layout, and presentation have significant impact on the ability of the user to identify and locate these features (Benbasat et al., 1986, John and Kieras 1996, Chuah et al., 1994). Moreover, many studies have shown that time to identify and locate information is reduced significantly with familiarity (John and Kieras 1996). Since engines display a great deal of information, it can be difficult to locate the desired feature on different engines, if not to know whether specific engines offer the feature at all! With repeated use, however, it becomes easier to locate various features, and develop reasonable expectations as to what would the results look like. In Appendix A, we plot the relative location of various (search and non-personalized) features on search engine interfaces. We also calculate the average area these features occupy on the screen. We observe that while the search boxes are roughly on the same location, and are highly visible on all the engines, the same is not true for other features. Some occupy a very small area on the screen. The engines are also highly disparate (in terms of location) in offering these features. As users start using the features repeatedly, they may find it easier to use the familiar features, which may lead to higher switching costs and hence more loyalty.

Finally, the personalized features are not only difficult to locate (perceptual task), they require the users to learn many rules (cognitive load). For example, to use an email, not only does a user have to locate and click it, but she has to go to a different interface that requires learning many rules about composing the message, sending it, etc. If the user *switches* to a different engine, she will have to learn many *new* rules making it difficult to transfer learning. Moreover there is a significant set-up cost when signing up for these features. For example, the user must select an id, password, and enter some personal information. In addition, switching to a different email facility would also require the user to communicate the new id to those who need to know the change. In sum, the switching cost for email is much higher than that of news or weather services. We expect the higher switching cost for personalized features would lead to higher loyalty. In addition, unlike search results which may lead to intra-session switching, there is little probability of intra-session switching for both personalized and non-personalized features.

3.3 User environment

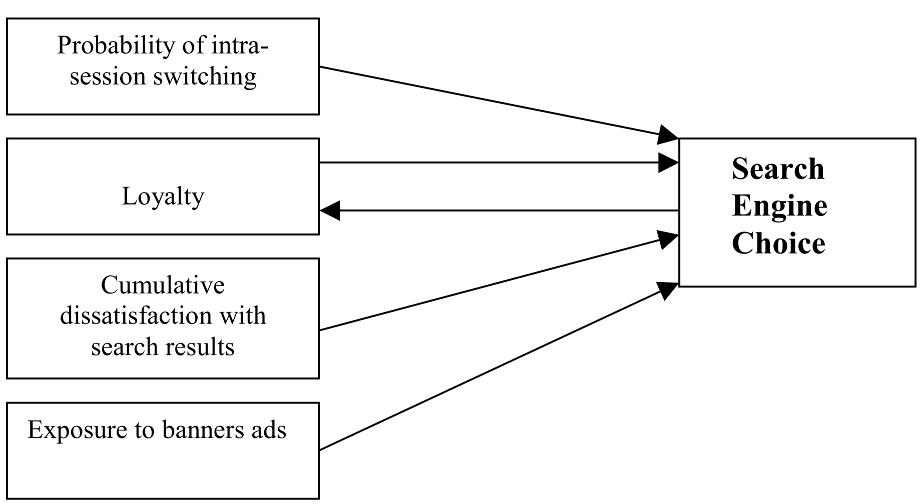
There is one aspect of the user environment, namely *Advertising Exposure* that may also affect the choice of a search engine. We consider only exposure to online advertisements, generally known as banner advertisements (Hoffman and Novak 1996). These advertisements have been shown to cause both attitudinal and behavioral changes (Briggs and Hollis 1997), although the click-through rates for the ads are reported to be low (Morgan Stanley 2000). Given the prevalence of advertising on the Internet, we have included these banner ad exposures in our choice model. We record whether the page visited immediately preceding a search engine choice contained an advertisement for that search engine. If it did, we consider the user to have been exposed to an advertisement. We believe advertisements that are seen immediately prior to the choice and have the ability to provide a direct click-through navigation to the sites, can have an impact on choice.

Figure 3 summarizes our model for search engine choice. We analyze this model in two steps. First, we examine user loyalty at the engine level as a whole without differentiating between search and non-search tasks. Next, we disaggregate loyalty for search task, and personalized and non-personalized features. We expect the strength of loyalty to vary across the three aspects of search engine use.

3.4 Formal hypotheses

Based on the forgoing discussions, we are now ready to lay out a series of formal hypotheses. Loyalty for search engines will arise due to the learning costs associated with becoming familiar with the interface and the format of search results of a particular engine. At the same time, dissatisfaction with the results will affect user choice behavior. In addition, those who actively use the personalized features of an engine are more likely to display loyalty to a particular search engine than those who do not. Likewise, non-personalized features use will enhance loyalty compared to the use of the search function alone. However, the effect of their use would be significantly less than that of personalized features. First, we hypothesize the effect of loyalty at the overall engine level.

Figure 3. A Model of Search Engine Choice



H1: *Past use of a particular search engine will have a significant positive effect on the probability that the same search engine will be chosen in the future.*

As we have explained, the use of multiple engines within a given search session can be due to unsatisfactory search results from the first engine chosen. So every time a user chooses an engine, there is a certain probability that an intra-session switch occurs. Formally:

H2: *There is a probability of intra-session switching for search tasks.*

We also argued that more intra-session switching would accumulate user dissatisfaction, and thus lower the probability of that engine being selected in the future. Formally:

H3: *Cumulative dissatisfaction with an engine will negatively impact the future choices of that engine.*

Next, we analyze loyalty at the task level. We expect that personalized features will be most effective in creating a loyal user base. Formally:

H4: *Loyalty for personalized features use will be significantly higher than the loyalty for non-personalized features or search features.*

We also explore the effects of advertising on engine choices. In particular, we examine the impact of banner advertising that occurs immediately prior to search engine choice. Thus:

H5: *The choice probability for a particular search engine will be increased if a user is exposed to a banner advertisement for that search engine immediately prior to the search engine choice.*

4. Model Estimation and Data Collection

This section is divided into three subsections. Section 4.1 discusses the base model, Section 4.2 describes the complete model, and Section 4.3 describes our data collection procedures.

4.1 Model

We combine the logit expression for choice probabilities with a linear utility model. The linear function contains time-varying attributes of given search engines as well as the response parameters associated with those attributes. For the base model examining user loyalty at the engine level as a whole, we write the deterministic component of utility as:

$$V_{ikt} = \beta_{ok} + \beta_1 Loyalty_{ikt} + \beta_2 Ad_{ikt} + \beta_3 CE_{ikt} + \beta_4 OS_{ikt} \tag{1}$$

β_{0k}	=	Brand constant for search engine *k*,
$Loyalty_{ikt}$	=	Loyalty of user *i* for search engine *k* at time *t*,
Ad_{ikt}	=	1 if user *i* was exposed to an online advertisement for search engine *k* at time *t* and 0 otherwise.
CE_{ikt}	=	Proportion of time user *i* switched away from engine *k* until time *t*.
OS_{ikt}	=	Occurrence of intra-session switching. If the engine chosen at time (t-1) was *k*, OS_{ikt} for engine k at time t is "1" because the user may make an intra-session switch. It is '0' for other engines. Note that β_4 indicates the probability of intra-session switching.

Hence, the probability of user *i* choosing an engine *k* at time *t* is:

$$\Pr ob(Y = k) = \frac{e^{V_{ikt} = \beta' X_{ikt}}}{\sum_{j=1}^{6} e^{\beta' X_{ijt}}} \tag{2}$$

where *X* is the vector of explanatory variables and β is the vector of respective coefficients to be estimated.

We note that in our sample, some user never exhibit intra-session switching behavior. We adopt the Mover-Stayer (MS) structure (Blueman, et al., 1955; Goodman 1961) in our model to account for this heterogeneity. We identify the users who never switch within a session as Stayers and the rest as Movers. Let *S* be the probability of a non-switcher and *L'* be the conditional likelihood function for a user. Let *Z* be an indicator variable such that *Z* = 0 for Stayer and *Z* = 1 for others. The likelihood function for a Mover is:

$$L = \sum_{Z=0}^{1} L'/Z.\Pr(Z) = (L'/Z=0).\Pr(Z=0) + (L'/Z=1).\Pr(Z=1) = L'(1-S) \quad (3)$$

The likelihood function for a Stayer is:

$$L = (L'/Z=0).\Pr(Z=0) + (L'/Z=1)\Pr(Z=1) = S + L'.(1-S) \quad (4)$$

Observe that $(L'/Z=0)=1$, since a Stayer always uses only one engine in a search session.

We construct the loyalty variable used in our utility model to examine the impact of past search engine choice on future choice probabilities. Guadagni and Little (1983) proposed an exponentially weighted average of previous purchases to capture the impact of past purchases, and termed it "brand loyalty" in their context. This measure for loyalty is very common in the literature (Papatla et. al., 1996; Gupta 1988), and fits the objective of this research well. Therefore, we adopt the G&L approach directly and write search engine loyalty as the recursive expression

$$Loyalty_{ikt} = \alpha.Loyalty_{ik(t-1)} + (1-\alpha).d_{ik(t-1)} \quad (5)$$

where:

$\alpha =$ carry-over constant

$d_{\mathrm{ik(t-1)}} =$ 1, if the user i selects search engine k at time (t-1) and 0 otherwise.

This particular conceptualization of loyalty is appealing because it directly measures the impact of previous choices on future choices. The carry-over constant α in our formulation and the loyalty coefficient β_1 in equation (1) cannot be jointly estimated using standard maximum likelihood techniques. Fader, Lattin, and Little (1992) proposed an estimation algorithm that allows joint estimation of both parameters. This is the estimation method we employ here.

4.2 Models estimated

In our analysis we estimate two models. For the first model, we estimate loyalty at the engine level data without differentiating among search tasks, personalized tasks, and non-personalized tasks. The utility function remains same as given in equation (1). For the second model, we differentiate between these three tasks. We create the loyalty variable, as in equation (5), but now it is specific to the three tasks. Depending on user choice, one of the loyalty variables will be active, and the other two will be "0."

So the utility function is:

$$V_{ikt} = \beta_{ok} + \beta_1 LoyaltyS_{ikt} + \beta_2 LoyaltyNS_{ikt} + \beta_3 LoyaltyPS_{ikt} + \beta_4 Ad_{ikt} + \beta_5 CE_{ikt} + \beta_6 POS_{ikt} \quad (6)$$

where:

$LoyaltyS_{ikt}$ = Loyalty for the search task of engine *k*, by user *i,* at time *t*. If the current choice is for search purpose then this variable is active, otherwise it is '0'.

$LoyaltyNS_{ikt}$ = Loyalty for the non-personalized features use of engine *k*, by user *i,* at time *t*. If the current choice is for non-personalized use then this variable is active, otherwise it is '0'.

$LoyaltyPS_{ikt}$ = Loyalty for the personalized features use of engine *k*, by user *i,* at time *t*. If the current choice is for personalized use then this variable is active, otherwise it is '0'.

After incorporating the Mover-Stayer structure, the likelihood function becomes:

$$L = \prod_{n=1}^{N}\left[\prod_{t=1}^{t_N}\left\{\left(\prod_{k=1}^{6}\left(Prob(Y=k)\right)^{d_{tk}} * (1-S)\right)^{I_n} * \left(\prod_{k=1}^{6}\left(Prob(Y=k)\right)^{d_{tk}} * (1-S) + S\right)^{1-I_n}\right\}\right] \quad (7)$$

where N is the number of users, t_N is the number of observations per user and $d_{tk} = 1$ if the engine k was chosen at observation *t*, and zero otherwise. $I_n = 0$ for Stayers and $I_n = 1$ for others.

4.3 Data on search engine use

The data used for this study comes from the HomeNet project (Kraut et. al. 1999) that tracked Internet use at home. The HomeNet project gave households a computer, modem, and an extra telephone line. Participants were also given three hours of training and free online support through a help newsgroup. HomeNet recorded a complete description of the Web sites visited by each user during each Internet session. Thus, for each user, we can determine which search engines she visited, what she did, and to what advertising she was exposed to prior to a search engine visit.

The most important characteristic of our data set is that it was collected in an unobtrusive and natural setting. Unlike many other studies on the Internet where researchers have used either survey data or laboratory experiments, which may bias the results (Mook 1983; Sears 1986), our data set has no such weakness. Participants access the Internet from their home environment, not a laboratory setting. The data collected contain actual choices rather than elicited preferences. These features provide a uniquely appropriate data environment for studying Internet portal usage behavior.

We collected Internet navigation data for 102 demographically diverse Internet users from June 1998 to June 1999. The data used in this analysis were assembled from detailed usage records captured at the server and took four person months of coding time. In the interest of keeping our subsequent estimation tractable, we captured search engine choices for the six most heavily visited (close to 90 percent) search engines during the study period, Yahoo!, Lycos, Excite, InfoSeek, Altavista, and HotBot. This data set contained 6,820 distinct search engine choices. The minimum number of data points for a user was seven, and the maximum number was 376. The mean number of search engine choices made by users over this period was nearly 67. The usual practice is to let the data initialize loyalty. So we used the observations from the month of June (n=499) to initialize loyalty and the rest of the data (N=6321) for model calibration. The aggregate choice frequencies in the data set were Yahoo! – 2,007, Lycos – 986, Excite – 1,490, InfoSeek – 1,056, HotBot – 337 and Altavista – 445. We also used five constants (β_{0k}) for Yahoo!, Lycos, Excite, HotBot, and InfoSeek, and dropped the constant for Altavista to achieve model identification.

Some important demographic characteristics of the users in this study are given in Table 1. Of the 102 subjects, a little more than half are women (53 percent) and almost three-quarters

(76 percent) are white. It can be seen from the table that the income and education of the subject pool have fairly wide ranges. It may also be noted that 59 percent of the subjects are adults (age>21).

To facilitate clear explanation, we adopted the following definitions in collecting and coding the data:

Session: We use the following rule to define a search session. Within a given log-on/log-off interval, a new session begins anytime more than 30 minutes elapse between search-engine visits. Our coding scheme accounts for the possibility that a user may have multiple occasions to use a search engine before exiting the Web. We later relax the assumption of the 30-minute interval and examine the sensitivity of our results to this scheme.

Table 1. Characteristics of the Subject Pool

Characteristic	Minimum	Maximum	Mean	Std. Dev.
Education (Yrs)	10	18	13.8	3.3
Income ($ '000)	17.5	85.0	52.1	23.6
Age (Yrs)	16	69.4	33.9	15.8

Search Engine Choice: From our data set we could easily distinguish the engine and the task performed by the user on that engine. We knew if the user performed a search and if so, what search term she entered. We also knew whether the user went to the engine to use mail, chat, or news services. It is common to observe a navigation pattern in which a user chooses a search engine, does some amount of query-based searching, visits some other sites, then returns to the engine (probably by using the "Back" button), and continues the search. The Web server used in this project captured this navigation pattern. Lest we overestimate loyalty, we considered only unique visits in a search session. Thus, within a given session, a visit to the search engine was considered a choice only the first time it was chosen. But if the user switched to another engine and then came back to the original engine and continued the search, we coded them as unique choices. For example, if a user visited Yahoo!, and Lycos, and then Yahoo! within half an hour to complete a search task, we coded these as unique choices. The navigation data were detailed enough so that we knew what search term users entered while using the search engine. Since we also know whether the user switched to another engine within the session, we can count the number of time users visited an engine and the number of times they switched to other engines due to dissatisfaction.

We also record whether or not the subjects use any of the personalized or non-personalized services. For example, we can verify whether the choices were made for e-mail, chat room, news etc., rather than search. For estimating our model with the utility function (6), we construct the appropriate loyalty variable by coding the choices for search, non-personalized, and personalized features. The following table displays the mean values (for n=6321) of our explanatory variables.

Table 2. Summary Statistics for the Model

Engine	*Loyalty*	*Ad*	*OS*	*CE*
Altavista	0.078	0.0005	0.0702	0.3365
Excite	0.2263	0.0710	0.2318	0.2812
HotBot	0.0639	0.0005	0.0527	0.2408
InfoSeek	0.1683	0.0090	0.1637	0.3181
Lycos	0.1495	0.0742	0.1531	0.3787
Yahoo!	0.3135	0.0035	0.3123	0.2473

5. Results and Discussion

We first estimate the baseline model (1) with engine level usage data. We test the hypothesis that loyalty plays a significant role in the choice of a search engine. We also want to test the effects of negative experience with the engine in future choices. The result of the estimation is reported in Table 3A. Note that the coefficients for a search engine (e.g., Excite

Constant) in this table show the intrinsic preference of users in choosing these engines relative to Altavista. A positive value indicates that they prefer that engine over Altavista.

Table 3A. Estimates for Engine level Choice

Parameters	Estimates (t-value)	Log likelihood	ρ^2
Excite Constant	0.474 (7.1)		
HotBot constant	-0.432 (4.8)		
InfoSeek Constant	0.430 (6.2)		
Lycos Constant	0.4467 (6.2)		
Yahoo! Constant	0.470 (6.8)		
S	0.100 (15.8)	-10346	
Loyalty	4.171 (42.3)	-8615	0.167
Ad	1.32 (16.7)	-8500	0.179
OS	-0.286 (6.7)	-8408	0.188
CE	-1.26 (12.0)	-8265	0.202

Our results clearly show that the past usage of the engine (loyalty) is a significant factor for future choices. Moreover, the sharp drop in the log likelihood clearly suggests that the loyalty variable explains a large portion of the variability in the model. Thus our results support hypothesis H1. A significant value of OS suggests that people exhibit intra-session switching due to dissatisfaction with the search results. The cumulative negative experience with the engine is also highly significant. In other words, users tend to use an engine less often if it fails to satisfy them often. Clearly, the engine's ability to satisfy the user with its results is an important predictor of the user's next choice. Both our hypotheses, H2 and H3, are supported by the results. The ad coefficient also appears to be significant. A prior exposure to an engine ad seems to increase the probability of choosing that engine. Note that the *Ad* variable indicates prior exposure, and not the click through rate. Finally the "S" value is also highly significant, justifying the use of the Mover-Stayer structure in our model. Incorporating M-S structure in our model also improves the likelihood.

For the second model, we create the loyalty variable at the task level as in equation (5). As we mentioned, learning the search task is simple and easy to transfer. Moreover, due to the very nature of the search process, users may exhibit intra-session switching. These factors collectively may not lead to a strong level of loyalty for search tasks. On the other hand, learning of personalized features is much harder to transfer to a different engine. Therefore, we expect that the use of personalized features will lead to a strong level of loyalty while non-search features and search tasks will have a comparatively weaker loyal consumer base. We needed at-least five observations to construct the meaningful loyalty variables using equation (5). Following this rule, there are 35 users who used personalized features for a total of 861 times in our data set, which translates to about 25 uses per user. Similarly, 65 users used the non-personalized features 1584 times, translating to about 25 uses per user for non-personalized services like news, weather etc. Note that a user may show varying loyalty to these features. We report the results in Table 3B.

Table 3B. Estimates for the Task Level Choice

Parameters	Estimates (t-value)	Log likelihood	ρ^2
Excite Constant	0.665 (9.9)		
Hot-Bot Constant	-0.152 (1.7)		
InfoSeek Constant	0.572 (8.1)		
Lycos Constant	0.600 (8.2)		
Yahoo! Constant	0.723 (10.5)		
S	0.098 (15.6)	-10346	
LoyaltyPS	6.020 (24.2)	-9237	0.107
LoyaltyNS	3.869 (21.5)	-8743	0.154
LoyaltyS	3.620 (34.4)	-8380	0.190
Ad	1.320 (16.8)	-8257	0.204
OS	-0.070 (1.7)	-8247	0.205
CE	-0.717 (6.9)	-8091	0.217

Interestingly, all three loyalty parameters come out to be significant. But parameter estimates, log-likelihood and ρ^2 numbers allow us to compare them further. Personalized features have the biggest β coefficient, followed by non-personalized and search features. Personalized features also have contributed most significantly to the likelihood. A sharp increase in the ρ^2 points out the significance of loyalty to personalized features. Non-search features come next, followed by search feature. Although loyalty for search is significant, its contribution to ρ^2 is not very high. A simple t-test shows that the coefficient of the personalized features is significantly higher than that of the non-personalized or search feature. Thus the results support hypothesis H4. Note that the coefficient of non-search features is not significantly different from that of search features.

More importantly, the results clearly point to the important phenomenon that the personalized features seem to be a major driving force behind the formation of loyalty. The use of non-personalized features or search feature does not lead to strong loyalty. The other parameters estimates are also fairly robust. Although their magnitudes change slightly, their signs do not change. Note that the value of OS is not significant at the 5 percent level anymore. But this is not surprising. Since the intra-session switching occurs only for the search feature, its relevance reduces at the task level loyalty estimation. As we have mentioned before, there is little chance of intra-session switching for personalized features.

Recalling that the G&L loyalty measure is a self-reinforcing construct where past purchases increase loyalty that in turn affects future choices, we can conclude that personalized feature use creates powerful incentives for users to patronize the same search engine. Using personalized features creates additional barriers to switching beyond those created by basic familiarity with the search engine interface and search output format. These increased switching barriers make the impact of past choices on future choice (loyalty) more potent. Moreover, it is also clear that search alone may not create enough loyalty for the engines, unless engines can offer high quality results. It is important for the engines to induce the users into using the non-search and, specifically personalized features. We tested various demographic variables (e.g., age, gender, income) for their impact on engine choices. Consistent with previous studies (e.g., Frank and Massy 1972), these effects are not significant.

Table 4A. Predictive Power of the Model

	Altavista	Excite	HotBot	InfoSeek	Lycos	Yahoo!
Probabilities	0.072	0.240	0.050	0.162	0.154	0.322
Actual N	445	1490	337	1056	986	2007
Predicted N	455	1520	314	1024	974	2032

We also test the predictive power of our model. At the sample mean given in Table 2, we present the predicted probabilities in Table 4A. Comparing the actual and predicted choice frequencies for each engine, we can see that the predictive power of our model is quite good. Finally, based on the loyalty values and intra-session switching properties, we know that the effect of loyalty will not be the same across all engines. In our sample, most of the personalized and non-personalized features were chosen for Yahoo! and Excite. Clearly, these two engines did an excellent job of converting the users into using their features. Given this, we would expect these two engines to have a much stronger loyal user base. To test our intuition, we calculated the elasticity of probabilities with respect to loyalty (Greene, 1993). Note that the "β" value of loyalty is not a slope parameter due to the non-linear logit model. We calculate the slope parameter for loyalty for six engines, and report the results in Table 4B. These estimates have the similar interpretation as the estimates in a linear regression. The results suggest that the marginal impact of loyalty on the probabilities of choosing Yahoo! or Excite is higher than any other engine. Clearly, in our sample, users show the strongest loyalty to Yahoo!, followed by Excite.

Table 4B. Elasticities of Probabilities with Respect to Loyalty

	Altavista	Excite	HotBot	InfoSeek	Lycos	Yahoo!
Elasticity	0.27	**0.76**	0.19	0.56	0.54	**0.90**

6. Conclusion

As Internet use and electronic commerce continue to expand, it is important that we understand how consumers interact with Internet portal sites. These sites hold powerful positions in the marketplace, guiding users to sites of potential interest. This research is an attempt to explore how Internet users choose search engines, a very common form of Internet portal site.

In this paper, we presented a conceptual model of search engine choice based on the literature of human-computer interaction. Our model is motivated by the notion that information goods such as search engines are a fundamentally different class of products than common household items. We hypothesized that the ability to learn various search engine features and the ease (or difficulty) of transferring this learning to other engines would determine loyalty in this context. We also explored why the user may exhibit differing levels of loyalty to search and non-search features of the engines. In addition, we posited that dissatisfaction with search results would negatively affect search engine choice.

Overall, we find that users develop loyalty for a given search engine. If they have used a particular search engine frequently in the past, they are much more likely to choose that search engine again in the future. The quality of the results is also a strong predictor of user choices. If the user is dissatisfied with the engine results, she might switch to another engine for the same query, and the probability that she will use the first engine in the future, diminishes. A first-mover advantage can take the firm only so far. A poor quality engine cannot hope to develop a loyal base.

We find that the search task alone does not develop strong loyalty. The loyalty becomes much stronger when the user starts using personalized features. This is a very important finding given the low barriers to entry in this industry as well as the low economic costs of switching engines. Search engines should focus on developing features that can be personalized. By providing such features they can improve the value proposition to users. Moreover, personalized features require additional learning, and thus increase the switching costs associated with patronizing another search engine. Our results very clearly point out that personalized features are very important in building a strong loyal base, although encouraging users to use them may be a challenging task. Whereas Yahoo! and Excite seem to have done a good job in converting search users in our sample into using non-search features, other engines have not.

As such, businesses using the Internet as a marketing channel often have used customizable features in an attempt to build customer loyalty. Amazon.com was first to introduce services that are customized to specific customer needs and tastes. Amazon began sending customers a list of suggested books and compact disks based on their past purchase behavior. In addition, Amazon invited customers to rate books and publish their own reviews on its site. Following Amazon's lead, other online retailers started similar programs (Tedeschi, 1999). NECX Direct, an Internet computer and electronics retailer, has been posting customer reviews on its Web site for the past two years. Our research suggests that these kinds of product development strategies are important for the long-term success of online retailers.

Our results are based on actual usage data from the server. Yet our subject pool is diverse and compares well with the Internet population of today (Rainie and Packel 2001). As mentioned previously, there are important advantages to using actual navigation data. This data, however, entails some restrictions. For example, we used a coding rule to infer the start of a new search session within a logon/logoff interval. To mitigate this restriction, we performed additional analysis indicating that our results did not change appreciably when we altered the interval to an hour.

The focus of this research is on the user choice behavior for search engines. We also studied the effects of using non-search features and dissatisfaction with the search results. While we have shown that loyalty is an important determinant of search engine choice, certainly other factors influence how an individual selects a search engine. It is important for future research to more fully characterize the decision process for these sites. The intra-session switching behavior and sampling of engines can be generalized to other product categories. We might expect to see similar behavior for choices for shop bots, newspapers sites, etc. We believe that more research is necessary to explore this issue further.

Further, our research directly suggests that it is not always optimal to make the features of a given Web site as transparent as possible. An axiom of human-computer interface design has been to improve the ease of learning. Interestingly, a profit-making objective may question this assumption. Indeed, requiring the user to learn a new interface and to supply personal identification data creates barriers to future switching behavior. However, if these features are made too difficult to use, people might never take the time to learn them. There would seem to be a delicate balance between the desire to make features easy to use and hence encourage trial and somewhat more demanding to use, and hence discourage future switching behavior. We hope future research will take our findings as a departure point for exploring how site features should be designed to leverage this trade-off.

References

Annual Report of Yahoo, 2000.

Austin D. (1997), "Ford Motor Company: Maximizing the Business Value of Web Technologies," *Harvard Business School Publishing*: Boston, MA, p. 9.

Ben-Akiva, and R. L. Steven (1985), "Discrete Choice Analysis: Theory and Application to Travel Demand," *MIT Press*, Cambridge-Mass.

Benbasat I., A. S. Dexter, and P. Todd (1986), "An experimental program investigating color-enhanced and graphical information presentation: an Integration of the findings," *Communication of the ACM*, 11, 1094–1105.

Blueman I., M. Kogan, and P.J. McCarthy (1955), "The Industrial Mobility of Labor as a Probability Process", vol. VI of *Cornell Studies of Industrial Labor Relations*, Cornell University, Ithaca, NY.

Briggs R., and N. Hollis (1997), "Is There Response Before Click-through," *Journal of Advertising Research*, 37(2), 33–45.

Bradlow E., and D. Schmittlein (1999), "The Little Engines that could: Modeling the Performance of World Wide Web Search Engines," *Marketing Science*, 19(1), 43–62.

Brynjolfsson E., and C.F. Kemerer, (1996), "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Software," *Management Science*, 42(12), 1627–47.

Carroll J.M. (1987), "Interfacing Thought: Cognitive Aspects of Human-computer Interaction," *MIT Press*: Cambridge, MA.

Carroll J. M., and M.B. Rosson (1987), "Paradox of the Active User," in the book *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction,* MIT Press: Cambridge, MA.

Chen Pei-Yu, and L. Hitt (2000), "Switching Cost and Brand Loyalty in Electronic Markets: Evidence from On-Line Retail Brokers*", Proceedings of 21st International Conference on Information Systems,* 134-44.

Chuah M.C., B. E. John, and J. Pane (1994), "Analyzing graphic and textual layouts with GOMS: results of a preliminary analysis," *Proceedings of the CHI '94 conference companion on Human factors in computing systems,* 323–25.

Davis S.A., and R.P. Bostrom (1993), "Training End Users: An Experimental Investigation of the Roles of the Computer Interface and Training Methods," *MIS Quarterly*, 17(1), 61–85.

Erdem T., and M.P. Keane (1996), "Decision making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1–20.

Fader P.S, J.M. Lattin, and J.D.C. Little (1992), "Estimating Nonlinear Parameters in the Multinomial Logit Model," *Marketing Science*, 11(4), 372–85.

Fortune (1999), "The Price on You Is What Drives Net Stocks," March 1, 217–218.

Frank R.E., W.F. Massy, and Y. Wind (1972), "Market Segmentation," *Prentice Hall Inc*: Englewood Cliffs, NJ.

Gentner D., and A.L. Stevens, (1983), "Mental Models," *Lawrence Erlbaum Associates*: Hillsdale, NJ.

Gilmore J.H., and B.J. Pine (2000), “Markets of One: Creating Customer-Unique Value through Mass Customization”, *Harvard Business School Press*: Cambridge, MA.

Goodman L.A. (1961), “Statistical Methods for Mover-Stayer Model”, *American Statistical Association J*., 56(December), 841-68.

Greene W.H. (1993), “Econometric Analysis,” *Macmillan Publishing Company*.

Guadagni P.M., and J.D.C. Little (1983), “A Logit model of Brand Choice Calibrated on Scanner Data,” Marketing Science, 2(3), 203–238.

Gupta S. (1988), “Impact of Sales Promotions on When, What and How Much to Buy,” *Journal of Marketing Research*, 25(4), 342–355.

Hawkins D.I., R.J. Best, and K.A. Coney (1986), “Consumer Behavior” (3ed), *Business Publications*: New York, NY.

Hoffman D.L., and T.P. Novak (1996), “A New Marketing Paradigm for Electronic Commerce,” *The Information Society,* 13 (Jan–Mar), 43–54.

Johnson-Laird P.N. (1983), “Mental Models,” *Harvard University Press*: Cambridge, MA.

Jeuland A.P. (1979), “Brand Choice Inertia as One Aspect of the Notion of Brand Loyalty,” *Management Science*, 25(7), 671–82.

John B., and D.E. Kieras (1996), “The GOMS family of user interface analysis techniques:

Comparison and Contrast,” *ACM Transactions on Human Computer Interactions*, 3(4), 320–351.

John B., and D.E. Kieras (1996), “Using GOMS for user interface design and evaluation: which technique?” *ACM Transactions on Human Computer Interactions*, 3(4), 287–319.

Kraut R., W. Scherlis, T. Mukhopadhyay, J. Manning, and S. Kiesler (1996), “The HomeNet Field Trial of Residential Internet Services,” *Communication of the ACM*, 39(12), 55–63.

Kraut R., T. Mukhopadhyay, J. Szczypula, S. Kiesler, and W. Scherlis (1999), “Information and communication: Alternative uses of the Internet in households,” *Information Systems Research*, 4(10), 287-303.

Lim K.H., L.M. Ward, and I. Benbasat (1997), “An Empirical Study of Computer System Learning: Comparison of Co-Discovery and Self-Discovery Methods,” *Information Systems Research,* 8(3), 254–72.

Lim K.H., I. Benbasat, and P. Todd (1996), “An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance,” *ACM Transactions on Computer-Human Interaction*, 3(1), 1–37.

Media-Metrix (2001), http://www.jmm.com/xp/jmm/press/mediaMetrixTop50.xml, October.

Money (1999), “Yahoo!’s Latest Deal Sways a Skeptic,” March 42.

Mobasher B., R. Cooley, and J. Srivastava (2000), “Automatic Personalization Based on Web Usage Mining”, *Communication of the ACM*, 43(8), 142-51.

Mook D.G. (1983), “In Defense of External Validity,” *American Psychologist*, 38 (April), 379–87.

Morgan-Stanley (2000), “The E-Marketing Report”, May.

Mukhopadhyay T., U. Rajan, and R. Telang (2001), “Competition Between Internet Search Engines,” GSIA working paper.

Mulvenna M.D., A. Sarabot, and A.G. Buchner (2000), “Personalization on the Net Using Web Mining”, *Communication of the ACM*, 43(8), 122-25.

Newell A., T.P. Moran, and S. K. Card (1983), “The Psychology of Human Computer Interaction”, *L. Erlbaum Associates*, Hillsdale, N.J.

Nault B.R., and A.S. Dexter (1995), “Added Value and Pricing with Information Technology,” *MIS Quarterly*, 19(4), 449–464.

Papatla P., and L. Krishnamurthi (1996), “Measuring the Dynamic Effects of Promotions on Brand Choice,” *Journal of Marketing Research*, 33(1), 20–35.

Petrison L.A., R.C. Blattberg, and P. Wang (1997), “Database Marketing: Past, Present, and Future,” *Journal of Direct Marketing*, 11(4), 109–125.

Rainie, L. and D. Packel (2001), “More Online, Doing More,” *The Pew Internet and American Life Project,* February.

Rao R.H., and B.D. Lynch (1995), “Hedonic Price Analysis of Workstation Attributes,” *Communication of the ACM*, 36(12), 95–102.

Sears D.O. (1986), “College Sophomores in the Laboratory: Influences of a Narrow Database on Social Psychologist’s View of Human Nature,” *Journal of Personality and Social Psychology*, 51(3), 515–30.

Schonberg E., T. Cofino, R. Hoch, M. Podlaseck, and S.L. Spraragen (2000), “Measuring success [personalised interactive services]”, *Communication of the ACM,* 43(8), 53-57.

Shugan S. (1980), “The Cost of Thinking,” *Journal of Consumer Research*, 7(2), 99–111.

Smith M.D., and E. Brynjolfsson (1999), “The great Equalizer? The Role of Shopbots in Electronic Markets”, *Presented in WISE, Charlotte, NC.*

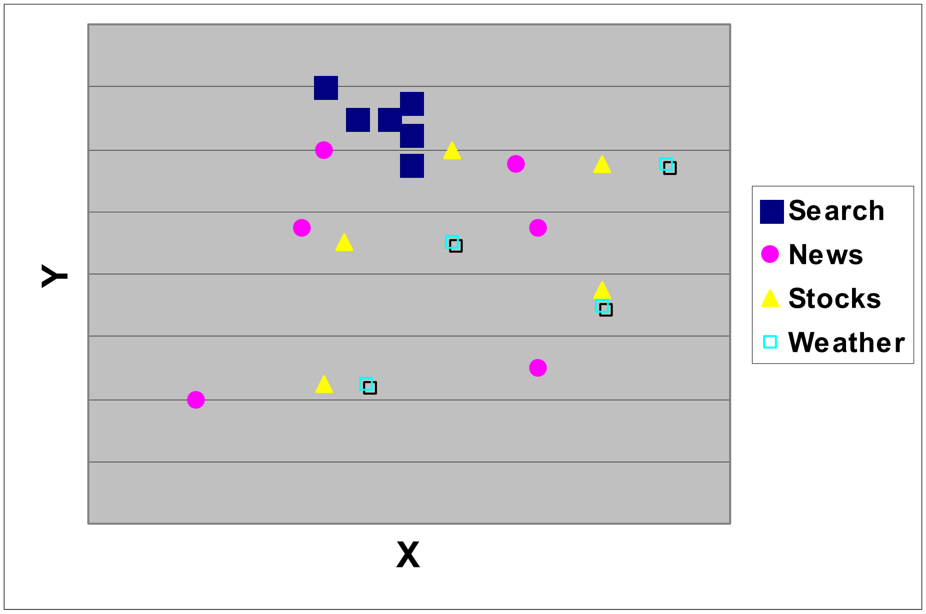
Staggers N., and A.F. Norcio (1993), “Mental Models: Concepts for Human-Computer Interaction Research,” *International Journal of Machine Studies*, 38, 587–605.

Tedeschi, B. (1999) “E-Commerce Report,” *New York Times*, September 6, p. C4.

Target Marketing (2000), “Top Ways Web Sites Are Discovered,” 23(4): 27, April.

Appendix-A

Relative location of search and three non-search features (News, Stocks and Weather) on different search engine interfaces.



The relative location of the search box is roughly the same on all engines. But the relative location of the other three non-search features is very different across the engines. There is little consistency across the engines when it comes to locating these features. Moreover, the average area of the search text box is close to (500x20) pixels compared to the average area for the non-search features that is (40x6) pixels. Most of the time, these features are in the form of a hyperlink. Given that there are on an average 570 hyperlinks on search engine interfaces, perceptually it is difficult to locate them, and then transfer this learning to other engines, which have different layouts. But repetition, by using the same engine again, can make the user efficient in locating these features.

The same is true for personalized features. They are not only difficult to locate, they essentially take the user to a completely different interface. Learning Yahoo! mail is not the same as learning Lycos mail. There are many new rules to be learned. Moreover, the user has to provide her id, password, and other personal information.