How Smart Are “Smart Banners”?  

Smart banners, or keyword-activated banners that are tied to internet user search keywords, are becoming increasingly popular with major search engines and their advertisers. Understanding how smart banners work best is a challenge facing the online advertising industry. This paper examines how specificity in the meaning of search keywords may affect the accuracy of banner matches. Through analysis of banner matches obtained from 12 major search engines, the authors found that as search keywords became more specific, search engines returned fewer exact banner matches and more general banner matches. Implications of these findings for search engines and their advertisers are discussed.

Advertising on the Internet has been growing steadily since the early days of the World Wide Web. The latest Internet Advertising Bureau/PricewaterhouseCoopers Online Advertising Report (October 3, 2000) confirmed that online advertising revenue has maintained its eighteen consecutive quarter of positive growth in the United States. In addition, the internet advertising industry broke the $2 billion mark in the second quarter of 2000, 127.3 percent over the comparative second quarter of 1999. With the rising number of online users and e-commerce activities, online advertising today has become increasingly popular for firms that intend to reach internet-savvy consumers. As the mainstream advertising community now knows, the internet as a direct marketing medium cannot only drive market share but also build brand (LeFurgy, 2000).

Currently, there are a variety of forms of online advertising, such as banner ads, sponsorships, and interstitials (Aronson, 1999). Among these, banner advertisements have been the predominant form so far. In the second quarter of 2000, banner ads accounted for 50 percent of the online advertising revenue, followed by 27 percent in sponsorships (Internet Advertising Bureau, 2000).

Banner ads are commonly placed on high-traffic websites, and they allow people to learn more about a particular product or service through hyperlinking, transferring the viewer directly to the company’s website for further information (English and Pearce, 1999). Studies have found that banner ads may significantly increase consumer awareness of online brands (Ipsos-ASI, 1999; Walker, 2000). Yet, banner ads have also been criticized as boring, and the average clickthrough rate has been declining (Cross, 1999). To address this concern, online media companies have tried to improve the appeal of banners through interactive banners that contain audio and video frames (Swibel, 1999), or to display targeted banners to more specific audience groups (Thompson, 1999).

Among targeted banners, keyword-activated banner ads (or smart banners) that pop up with user input keywords are probably the most frequently used, especially by major search engines. For instance, auto-buy-tel.com could buy a “smart banner” that is prompted only by the keyword “auto” typed by internet users. Pioneered by Infoseek in 1995, smart banners have been growing in popularity; Strategis Group (1999) estimated that 83 percent of Yahoo! searches and 80 percent of Lycos searches yield related banner advertisements. Most of the search engines have adopted this practice.

A major promise of smart banners is that they
SMART BANNERS

can increase the effectiveness of online advertising because companies can use them to pursue particular audiences who share common and specific interests, e.g., consumers who type the "auto" keyword to search for information about automobiles. When internet users type a particular search keyword (e.g., auto), they are more likely to notice a pop-up banner (e.g., auto-by-tel.com) that is closely related to their demonstrated interest at that time. The enhanced interest is then likely to motivate the users to click the banner to find more information. In fact, this is the exact rationale behind major search engines charging higher advertising rates for "smart banners"—about $50 per thousand impressions compared to $20 for average banners (Frauenfelder, 1999).

Despite the prevalence of smart banner ad use by major search engines and their clients, no empirical studies have actually examined the performance of the so-called smart banners. No clear performance criterion has been proposed to define a "smart" banner and assess its degree of "smartness." In addition, the lack of comparative-performance data for "smart banners" across major search engines makes it difficult for potential clients to assess the ability of each search engine to deliver targeted messages tied to keyword searches. Further, for major online media companies (e.g., search engines), a common performance yardstick could also be valuable for comparing performance, benchmarking, and quality improvement purposes. Toward that goal, search engines need to identify critical factors that may affect the performance of their "smart banners" to improve accordingly.

Both supply-and-demand-side factors may decide whether a search engine can produce exact banner matches for internet search terms. Demand side factors may include how internet users choose to use different levels of search keywords and how many online advertisers are interested in purchasing banners for the particular keywords. Supply-side factors may include how many banner ad copies are available to be matched with keywords and how many search keywords a search engine plans to sell for targeted banners.

This article endeavors to evaluate the impact of a particular demand-side factor on banner matches, i.e., how does the specificity of consumer search keywords (e.g., cars vs. sporty cars) influence whether search engines can produce exact banners that correspond to keywords precisely (i.e., displaying a banner for auto-by-tel.com vs. a banner for greed4speed.com)? Along those lines, a number of pertinent questions await answers. For instance, are some search engines doing a better job than others at delivering banner ads that are closely tied to user searches? How effective are search engines in delivering exact banner matches to user searches? How does specificity in the meaning of keywords affect the accuracy of banner matches? In the absence of exact matches, can the search engine deliver close matches for user keyword searches? Do natural language search engines (e.g., Ask Jeeves) do a better job of displaying keyword banners through the correct interpretation of search keywords?

This study is designed to fill a void in the research with empirical evidence about the effectiveness of smart banners and key success factors. Its central objective is to examine major search engines and rate their ability to produce relevant banner advertisements based on selected keywords. It also aims to investigate how specificity in the meaning of search keywords affects the accuracy of banner matches. The key hypothesis of the study is that as keywords become more specific, the number of exact banner matches will decrease; banners with broad object meanings will become more common; and banners with more specific meanings will become less common. To check the validity of this hypothesis, it was examined for 12 individual search engines as well as for all of them combined. In addition, we checked whether the popularity of search engines or keywords might affect the test results.

We first compiled a list of 115 main-level search keywords using the root or first-level categories from a variety of directory-type search engines (e.g., Yahoo, Excite, and Lycos). We then expanded the list to build more specific search phrases using the subordinate levels of search categories. The final list of 345 search keywords or phrases with different degrees of specificity in the meaning of the search keyword were then entered into 12 major search engines to find the resulting banners. We rated the displayed banners corresponding to those keywords for relevance and degree of match to the keywords. We found that as internet users' search keywords become more specific, there were fewer exact matching banners. The findings of the study provided firsthand empirical evaluations of smart ad performance across major search engines. The study also identified the keyword specificity as a critical factor affecting smart ad performance. As such, the study results will help search engines improve their smart ad performance and help online advertising companies make better use of smart banner ads.

CLASSIFYING INTERNET SEARCH KEYWORDS

On any typical day, a major search engine may serve millions of online users with diverging search interests (MediaMetrix, 2000). Obviously, for companies that sell differentiated products/services, delivering the right messages to the right cus-
delivering the right messages to the right customer groups is an important advertising goal.

The classification scheme for internet users’ search keywords proposed in this research is consistent with human language relational models discussed in the semantics literature (e.g., Hatch and Brown, 1995; Saeed, 1997). According to these models, relationships between search words described in the “broad-moderate-narrow” scheme can be regarded as hyponym relations in which the “moderate” keyword such as “air travel” is a hyponym of “broad” keyword such as “travel.” Further, a narrow keyword such as “discount air travel” is a hyponym of “moderate” keyword such as “air travel.”

As pointed out by linguistics researchers (e.g., Kastovsky, 1990), hyponym relations are important for semantic descriptions and memories of words. In the context of internet user searches, we believe that the proposed “broad-moderate-narrow” classification scheme based on hyponym relations borrowed from the semantics field is appropriate for the first step in a study about keyword banners; this scheme will assist in understanding the different types of search keywords internet users may select and the relationships among them.

The three types of internet search keywords or phrases we propose have indeed been discovered in internet users’ actual search terms. Data compiled by a number of search keyword monitoring agencies (e.g., WordSpot, KeywordCity, and search voyeur sites provided by search engines such as Excite) have confirmed that while single keyword searches for “broad” categories are popular, a significant number of internet users are also refining their searches using “moderate” or “narrow” keywords. For instance, KeywordCity, using actual search terms collected from GoTo.com during May, June, and July of 2000, reported that about 40 percent of hotel-related searches used the broad keyword “hotels”—while “moderate” or two-word keyword phrase searches (e.g., Las Vegas Hotel, discount hotel) accounted for 35 percent of the total “hotel” related searches. “Narrow” keywords, or three-word search phrases such as “pet friendly hotel,” accounted for 20 percent of the searches. The same pattern was also observed for other common search categories such as “travel” or “vacations.” Further, findings based on data released by Search Keyword Monitoring services are consistent with reports from other sources. For instance, a recent NPD New Media Services (2000) recent survey of 33,000 internet users picked at random during the first quarter of 2000 found out that nearly 45 percent of web users searched using multiple keywords, a far more popular method among respondents than using one keyword (28.6 percent). Thus, we believe that the proposed classification scheme is useful as internet users were found using simple search keywords as well as more complex search phrases.

MATCHING KEYWORDS WITH SMART BANNERS

Studies have found that internet searchers rarely look at more than 30 search results (or the first several pages) returned by search engines (Greensberg, 2000). This implies that the rectangular screen real estate allocated for a resulting banner may be very important in capturing searchers’ attention. If the banner is “smart,” and highly relevant to the user’s search, then it is likely to be granted the same or even...
SMART BANNERS

greater importance than the first few text links because banners can convey both text messages and rich images. The matching of keywords and banners can be commonly achieved either through words, or images, or both in the banners. For instance, etoys.com may have banners with words or images that highlight girls’ toys (e.g., Barbie dolls), or boys’ toys (e.g., truck models).

Next, we analyze a spectrum of possible banner match scenarios from the user’s perspective; each of them represents a different level of “smartness.” If we put the five different types of matches on a continuum of “smartness” for keyword-activated banners, then we define “exact matches” as “very smart,” “upward or downward partial matches” as “smart,” “relevant matches” as “marginally smart,” and “irrelevant” banners as “not smart at all.”

1. Exact match. This is certainly the most ideal scenario for both the internet searcher and banner advertiser. An example is an appearance of a “buycomp. com” banner for the keyword “computer.” Obviously, the banner message clearly matches the interest of the searcher at that point.

2. Upward partial match. In this scenario, a relatively specific keyword produces a banner that encompasses more than the search goal. For example, when a user types “toys for young girls,” a banner for the general store etoys.com is displayed. Or when users type a specific brand (e.g., IBM) and they see a banner for a vendor that carries the brand (e.g., cdw.com). Consequently, even if users are looking for something more specific, they may be given a related yet broader banner. Only “moderate” and “narrow” types of keywords can prompt “upward partial matches.”

3. Downward partial match. In this scenario, a relatively broad keyword produces a banner that encompasses more-refined domain than the search goal. For example, when users type “personal finance,” they are shown a banner for etrade.com, which deals primarily with online stock trading. Here, users may see a relevant banner but it may contain narrower offerings than those being asked for. According to the keyword classification scheme, only “broad” and “moderate” keywords can prompt “downward partial matches.”

4. Related matches. This category denotes banners that are not directly related to the keywords but may have some connections with the keywords so they may not be completely irrelevant from the users’ perspectives. An example is the appearance of a banner for “monster.com,” an online job placement website, when a user searches for “career training.” Users may believe that “career training” is certainly relevant for finding a job so they may not disregard the banner completely. Thus, an online media company may find displaying related matches for keywords has some utility. All three types of keywords are equally likely to encounter this type of match.

5. Irrelevant. This category denotes categories that are unrelated to the keyword search. An example could be a banner for WebMD (a medical information portal) displayed when a user searches for “heavy-duty mountain bikes.” From the searcher’s point of view, the displayed banner is certainly not related and not worthy of attention.

Next, search engines may fail to fully utilize different versions of a client’s banner to conform to different levels of users’ searches. For example, if an online toy retailer has developed a variety of banners to suit the variations in customer search keywords, then the search engine can match different banners with different
search keywords: a general store banner for the search keyword “toys”; a store banner with images of Barbie dolls for the search keyword “toys for girls”; and a store banner with images of truck models for the search keyword “toys for boys.” Even though a few companies are pursuing this route in search-engine banner advertising (Kuchinskas, 2000), casual observations of online banners suggest that this practice is hardly the norm. Last, the search engines’ algorithms may not be intelligent enough to correctly interpret complicated searches. For instance, a search engine may return a banner for an online discount retailer for the search keyword “discount travel.”

The above discussions about factors affecting types of banner matches provide motivations for the following research hypotheses:

H1: As search keywords become narrower, exact banner matches will be fewer.

It will be more difficult for search engines to produce exact matches when the keywords become too specific. This is true because, as keywords narrow, there will be less supply of more specific keyword banners and less demand for buying specific search phrases.

H2: As keywords become narrower, more upward partial matches are likely.

If a search engine cannot produce exact banner matches (e.g., discount travel to Europe) because either the keyword is not for sale or is not bought, it is likely to produce banners that are as relevant as possible by attempting to match at least one of the recognized keywords (e.g., travel). Usually, this means displaying the banner for a popular “broad” keyword (Gursky, 2000; Heller, 2000). This approach is more sensible than displaying a random run-of-site banner because it enhances the advertiser’s reach to users who are pursuing similar search concepts. Since this approach augments both the advertiser’s and the searcher’s benefits by displaying relevant banners, we expect major search engines to prefer this approach to a run-of-site banner.

H3: As keywords become narrower, fewer downward partial matches are likely.

The rationale here is similar to that given for H1. If a search engine encounters difficulties in supplying exact banner
matches as keywords become narrower, then it will most likely fail to supply banners that delineate a narrower business boundary than that implied by the search concept. For example, if a search engine could not display an exact banner match for the search keyword “toys for toddlers” because of its limited keyword banner ad buys, then it must be even tougher for this search engine to find and display an even narrower banner about learning software for preschoolers. The first three hypotheses are illustrated in Figure 2. We expect these three hypotheses to hold true for each individual search engine that pursues keyword banners.

H4: Natural language search engines will behave in a way similar to other search engines by producing keyword banners as search keywords become narrower.

Natural language search engines (e.g., Ask Jeeves), claim to understand complex search expressions or search questions. While this capability is certainly desirable for finding exact information on the internet, we posit that the types of ad banner matches have more to do with relevant banner ad inventories and ad-keyword matching capabilities. In addition, other search engines should also be able to correctly interpret internet search keywords in most instances (Sherman, 2000). Therefore, we expect search engines such as Ask Jeeves to conform to hypotheses H1 to H3.

H5: Top e-commerce keywords are likely to generate more exact matches, and non-top terms are likely to generate more irrelevant matches. H1 to H3 apply to both types of keywords.

Given the sheer number of online bookstores (several hundred listed in the Yahoo! directory), we expect to see more exact matches for “books” than for “motor oil” as the demand for the former keyword is likely to be larger. Indeed, the demand for smart banners that are tied to top e-commerce keywords is so high that search engines (e.g., GoTo.com) may charge almost 10 times more for them than for “non e-commerce” keywords (Thompson, 1999). “Hot” e-commerce keywords, such as “books” or “travel,” are thus likely to have more buyers and greater inventories, making them amenable to producing more exact matches.

Further, as the underlying rationale behind H1 to H3 should hold for any keyword, we posit that H1 to H3 will still apply to both top and non-top e-commerce terms.

H6: Top search engines are likely to generate more exact matches, and non-top search engines are likely to generate more irrelevant matches. H1 to H3 apply to both types of search engines.

There are 24 major search engines that are commonly known to internet users (SearchEngineWatch, 2000). Among those, only a few of the more popular ones (e.g., Yahoo!, MSN) can attract the most advertisers and revenue because of their high traffic numbers (Forrester Research, 2000). Further, popular search engines are more likely to deploy keyword banners (Thompson, 1999) as they probably can afford to commit the financial and technological resources needed to implement the sophisticated matching between varied keywords and diverse banners. Thus, top search engines should stand a better chance than non-top search engines in producing more exact matches and fewer

![Figure 2](https://example.com/figure2.png) Changes in the Number of Different Types of Matches as Keywords Become Narrower

36 JOURNAL OF ADVERTISING RESEARCH July • August 2001
irrelevant matches. Further, as the underlying rationale behind H1 to H3 should hold for any search engine that uses keyword banners, we posit that H1 to H3 will hold for both top and non-top search engines.

Finally, there is no compelling reason to believe that as keywords become narrower, the "related" or "irrelevant" categories will either increase or decrease. This is certainly logical if the search engine chooses to display a random "run-of-site" banner in the case of nonrecognition of a keyword. On the other hand, if a search engine can recognize the "broad" search keyword and attempts to display a "related" banner, then it should at least recognize the same keyword in the narrower keyword phrases so as to display "related" banners for the narrower search terms. In both scenarios, the search engine is unlikely to display either significantly more or significantly less "irrelevant" or "related" banners as keywords become narrower. Consequently, no directional hypotheses are given concerning the impact of keyword specificity on the last two type of banner matches.

METHODOLOGY
A sampling field study was designed to provide insights into the performance of "smart banners" and to test our research hypotheses.

Sampling of keywords
Theoretically, any combination of words in a dictionary may be used as search keywords or combined to build search phrases, so there does not exist a comprehensive sampling frame of internet search keywords that can be used for our study purpose. Hence, we adopted a judgmental sampling approach (McDaniel, 1999) in this study.

We followed the semantics principles regarding the construction of hyponyms (e.g., Saeed, 1997) in building a sample of "broad," "moderate," and "narrow" search terms. "Broad" keywords were picked from the root categories or first-level categories in directory-type search engines (e.g., Yahoo, Excite, and Lycos); examples include "travel," "computers," and "furniture." "Moderate" keywords were compiled either from the next level of search categories (e.g., air travel) or by adding modifiers to the "broad" keywords; examples include "discount travel," "laptop computer," and "sofa." Finally, "narrow" keywords were either chosen from even lower levels of search categories listed in a directory of search engines or developed by adding more modifiers to the "moderate" keywords; examples include "discount travel to Europe," "laptop computer accessories," and "leather sofa."

In compiling this judgmental sample of keywords, a number of factors were carefully considered. First, for-profit organizations comprise the majority of online advertisers so we chose not to include keywords for non-profit organizations (e.g., government) or non-commercial keywords (e.g., science, religion). Second, keyword categories that are unlikely to be of general consumer interest (e.g., wholesale, generators) were not selected. Third, controversial (e.g., pornographic) terms were not selected. Finally, the names of people (e.g., Britney Spears) or companies’ URLs (e.g., www.ge.com) were not chosen as search terms. In general, the search expressions chosen encompassed a representative sample of terms that ordinary internet users may use to search for product or service information and for which companies are likely to buy online banners.

Since we also wanted to compare the effectiveness of search engines at responding to "top" as well as "non-top" e-commerce terms, we assigned the compiled keywords to one of those categories. We defined a term "top" based on information compiled from a variety of trusted online research agencies. For instance, The Standard (2000) identified books, videos, music, and software as accounting for 23.3 percent of e-commerce sales in January 2000. The National Retail Federation/Forrester Research (2000) listed the following categories as the Top 10 online retail categories in February and March 2000: software, music, health and beauty, food and beverage, books, apparel, consumer electronics, hotel reservations, computer hardware, and air tickets. The 10th GVU World Wide Web user (2000) survey shows that internet users frequently search for information on the following subjects: news, reference material, and financial information.

A selected portion of the sample keywords used in this study is given in Table 1 (on the following page). The final sample contains 115 broad categories with a total of 345 individual keywords at three different levels.

Data collection
Data collection was conducted through the Namestake Banner Check service at www.namestake.com, a division of Thompson & Thompson. The website can simultaneously display banners from 13 different search engines (2 of which were dropped from the study because they produced very few banners) by typing a single keyword or phrase. This convenient feature ensures that the researcher can compare the banner matches for 11 different search engines at once. The 11 search engines studied are: Altavista, Excite, Hotbot, Infoseek, Lycos, NetFind, Netscape, Snap, Search, Webcrawler, and Yahoo. Additional data of banner matches from ask.com using the same set of keywords were also collected. Since all search-engine banners (except those from Ask
TABLE 1
Sample of Keywords Used in the Study

<table>
<thead>
<tr>
<th>Type of Keyword</th>
<th>Keyword</th>
<th>Top E-Commerce Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad</td>
<td>Books</td>
<td>Yes</td>
</tr>
<tr>
<td>Moderate</td>
<td>Kids' books</td>
<td>Yes</td>
</tr>
<tr>
<td>Narrow</td>
<td>Cartoon books for kids</td>
<td>Yes</td>
</tr>
<tr>
<td>Broad</td>
<td>Toys</td>
<td>Yes</td>
</tr>
<tr>
<td>Moderate</td>
<td>Girls' toys</td>
<td>Yes</td>
</tr>
<tr>
<td>Narrow</td>
<td>Toys for toddler girls</td>
<td>Yes</td>
</tr>
<tr>
<td>Broad</td>
<td>Furniture</td>
<td>No</td>
</tr>
<tr>
<td>Moderate</td>
<td>Sofa</td>
<td>No</td>
</tr>
<tr>
<td>Narrow</td>
<td>Leather sofa</td>
<td>No</td>
</tr>
<tr>
<td>Broad</td>
<td>Musical instruments</td>
<td>No</td>
</tr>
<tr>
<td>Moderate</td>
<td>Electronic musical instruments</td>
<td>No</td>
</tr>
<tr>
<td>Narrow</td>
<td>Electronic musical instruments for kids</td>
<td>No</td>
</tr>
</tbody>
</table>

Jeeves) are displayed on the same page, research time is reduced, and banners displayed at different search engines can be recorded immediately.

To ensure that the criterion used to judge the degree of match was consistent, the authors first practiced identifying types of matches and extensively discussed results associated with 15 categories. The remaining categories were checked individually and cross-validated. The data collection period was limited to a month (March 2000), so it is reasonable to assume that there were no major changes in internet user profiles or search engine target banner policies.

RESULTS
In this section, we present our analysis of the sample data. Descriptive results are presented first followed by hypothesis testing results.

The first goal of the study is to see which search engine produced the most banner matches. Not all of the search engines returned banners for every keyword (e.g., because the server was busy); the average is 315 banners for 345 keywords. As a result, we chose to look at the relative percentages of different types of banner matches obtained for each search engine. Figure 3 provides an example of the relative performances of the search engines being studied.

As Figure 3 shows, these search engines returned a surprisingly large percentage...
SMART BANNERS

of irrelevant banner matches. Search.com produced the largest portion of irrelevant banners, about 93 percent. AltaVista did best in producing exact matches (26 percent), followed by Excite (22 percent) and NetFind (22 percent). Regardless of the search engines, upward partial matches are more frequently shown than downward partial matches, indicating that search engines tend to display banners that cover broader domains if user keywords are too specific to handle. Hence, providing downward partial matches proved to be a daunting challenge for all search engines.

The number of different types of matches varied considerably among the 12 search engines. AltaVista had 65 percent pertinent matches, including “exact,” “upward partial,” “downward partial,” and “relevant.” Netscape and Excite came second and third, with 63 percent and 59 percent of pertinent matches, respectively. Interestingly, the so-called natural language search engine Ask Jeeves demonstrated a mediocre performance in returning banner ads that are tied to user search engines; only 53 percent of the banners displayed were pertinent. Overall, we conclude that major search engines still have plenty of room for improvement in producing banner ads that closely match internet user search keywords.

**Narrowness of keywords and types of banner matches**

Here we present empirical testing of the major study hypothesis as embodied in Figure 2. After aggregating the results from all 12 search engines, we saw several ways the matches change as the keyword becomes narrower (see Table 2). We found that, as predicted, the number of exact matches decreased from 301, to 182, to 110 as the type of keyword shifted from broad, to moderate, to narrow. At the same time, upward partial matches increased from 20, to 153, to 222. Downward partial matches dropped from 137, to 52, to 7 as keywords became narrower. As for “related” matches, the percentage changes were not dramatic and uni-directional as compared to the previous three types of matches. The number of “related” matches changed from 106, to 132, to 144, respectively. Finally, the number of “irrelevant” banners displayed did not vary much as keywords become narrower, from 495, to 570, to 597.

To further demonstrate the significance of the above findings, a Chi-square analysis was conducted to test the null hypothesis that there is no significant relationship between the type of keywords and the particular type of banner matches. The Chi-square analysis result (p < 0.001) suggested that our null hypothesis should be rejected. This indicates a definite association between the narrowness of a keyword and the type of banner matches, especially for the first three types of matches: exact, upward partial, and downward partial. For instance, when a keyword changes from broad to moderate, the number of exact matches drops by nearly 40 percent. If a narrow keyword is used, then the drop is another 40 percent. On the other hand, the correlation between the narrowing of keywords and “related” or “irrelevant” banners is not strong.

The results discussed thus far are based on aggregated data from all 12 search engines in the sample. We decided to check whether the results might be different if the above hypothesis was tested on the 12 search engines separately. Thus, 12 additional Chi-square independence tests were conducted for each of the 12 search engines. With the exception of Search.com (whose large percentage of “irrelevant” matches rendered the Chi-square test invalid), all search engines displayed the same pattern and significant results in the Chi-square tests. Therefore, we conclude that H1 to H3 were strongly supported for all the 12 search engines combined as well as for each individual one (except search.com).

Lastly, the performance of Ask Jeeves, the natural language search engine, displayed the same trend shown by other search engines. So, H4 was also supported by the empirical analysis. Hence, we confirmed that there is no inherent advantage to “natural language” search engines; they provided no more precise banners to match more specific keywords, at least based on this sample of data.

**Table 2**

<table>
<thead>
<tr>
<th>Type of Keyword</th>
<th>Exact</th>
<th>Upward Partial</th>
<th>Downward Partial</th>
<th>Related</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad</td>
<td>301</td>
<td>20</td>
<td>137</td>
<td>106</td>
<td>495</td>
</tr>
<tr>
<td>Moderate</td>
<td>182</td>
<td>153</td>
<td>52</td>
<td>132</td>
<td>570</td>
</tr>
<tr>
<td>Narrow</td>
<td>110</td>
<td>222</td>
<td>7</td>
<td>144</td>
<td>597</td>
</tr>
</tbody>
</table>

Do top-e-commerce keywords produce more relevant banners?

We also examined the impact of a “top” keyword on the accuracy of banner ad matches. The results are shown on the next page in Figure 4. We found that top keywords were better at producing exact, upward partial, and downward partial matches. They also produced fewer irrel-
relevant matches and about the same number of related matches. In fact, exact matches were found for about 20 percent of the top e-commerce keywords, compared with 17 percent for non-top e-commerce keywords. In addition, irrelevant matches were discovered for 48 percent of the top e-commerce keywords compared with 54 percent for non-top keywords. While top e-commerce keywords seemed to perform slightly better, the differences were not dramatic.

The result surprised us; we expected that search engines would at least produce many more relevant banners for "top" e-commerce keywords (e.g., travel). This finding suggests that online media companies should focus on improving banner ad matches for "top" e-commerce keywords that are frequently used by internet users and sought after by advertisers. Thus, building a large inventory of banners that are tied to "top" e-commerce keywords should be the first step toward building "smarter" banners.

To check whether H1 to H3 hold for "top" and "non-top" e-commerce keywords, we constructed a table (not shown here) that is similar to Table 2 except that only the "top" e-commerce keywords were used. We found the same pattern associated with changes in how many banner matches were made in each category as keywords become narrower. Further, the Chi-square test was still significant, supporting the same conclusion that was reached based on all types of keywords. The same procedure was then repeated for non-top e-commerce keyword, and the same pattern of matches was found.

Overall, H5 was supported by the empirical data. Thus, we conclude that while "top" e-commerce keywords may produce more exact banner matches than non-top keywords, both types demonstrated similar behaviors in displaying relevant banners as the keywords narrowed.

Figure 5 Impact of Search Engine Popularity on Types of Matches

We also examined whether the popularity of search engines affects how they produce keyword banner matches. Using MediaMetrix (2000) ratings, we classified AltaVista, Excite, Infoseek, Lycos, Netscape, and Yahoo as "top search engines," and Hotbot, NetFind, Snap, Search, WebCrawler, and Ask Jeeves as "not-top search engines."

We found that top search engines performed consistently better in the four pertinent categories (i.e., exact, upward partial, downward partial, and related) (see Figure 5). They also returned fewer irrelevant matches than the "non-top" search engines. Our finding here is consistent with H6, which stated that top search engines will better support smart banners because they are likely to have the technological expertise needed for complex banner-keyword matching and possibly to possess larger banner ad inventories.

To test whether H1 to H3 applied to the two "popular" and "less-popular" search engine groups, we ran a Chi-square analy-
... the more accurate the banner ads are for the corresponding keywords, the more likely the advertiser will see the value of its online advertising dollars.

This study was undertaken to evaluate the effectiveness and success factors of this important online advertising tool. Few empirical studies have actually evaluated how “smart” those smart banners are or the factors that may determine whether there will be exact banner matches for different types of search keywords. To solve those puzzles, our study first proposed a semantics-based framework for classifying internet search keywords. Then we suggested a structure that categorizes different degrees of accuracy in matching banner ads to user keyword searches. We also hypothesized that the specificity of search keywords will influence the type of matches a user receives. We expected this fundamental hypothesis to be invariant regardless of the popularity of search engine or search keywords. We also speculated that popular search engines were likely to generate more exact banner matches than less popular ones, especially for top e-commerce keywords.

All hypotheses were tested using empirical data and were strongly supported. Finally, this study also compared the relative performance of major search engines at producing “smart banners.” Overall, results of this study will contribute to a deeper understanding of “smart banners” and how to improve their performance. Search engines can use the study results for benchmarking or quality improvement purposes.

**Implications for search engines**

Advertising represents a significant portion of search engine revenues, e.g., 70 percent for Yahoo! (Katz, 2000), and banner ads are a major online advertising tool (Internet Advertising Bureau, 2000). Obviously, making banner ad performance accountable is a big challenge for search engines. The implementation of smart banners has the potential to increase banner ad effectiveness by displaying them at the right point of a consumer’s information search process.

Search engines need to work closely with advertisers and their agencies in developing, testing, and managing effective keyword banners. A possible future direction for search engines is to develop advanced advertising administering technology to handle adaptive banners that can incorporate real-time user search keywords into a pre-built banner template for the advertiser. For instance, a travel site that uses its brand logo as the base banner template can embed appropriate user search banner keywords (e.g., budget travel for students) into the base banner. The application of adaptive banners may reduce the reliance on pre-built banner inventories and significantly enhance the relevancy of banners tied to internet search keywords.

On the other hand, when the most desirable exact matches are not economically feasible, (e.g., because of high development cost in advertising administering technology), then search engines should still make sure that internet users at least see partially matched banners. For instance, if a user types “Pokemon trading cards,” the search engine can return an eToys banner through an upward partial match. Such partially matched banners may still be of interest to internet users or provide a new search direction for them.

**Implications for advertisers**

For advertisers, the more accurate the banner ads are for the corresponding keywords, the more likely the advertiser will see the value of its online advertising dol-
SMART BANNERS

An advertiser can contribute to the enhancement of banner match accuracy by supplying varied forms of the same company banner to the search engine. Fortunately, the industry seems to start noticing the importance of varied banner versions. For instance, Organic—an interactive agency based in San Francisco—worked with an online music retailer that was running a few different banners against keywords on search engines. The agency developed 250 different banners that were tied to narrower search terms and the advertising campaign produced a 600 percent increase in revenue over one year because of more enthusiastic user response to the banners (Kuchinskas, 2000).

In addition, an advertiser that desires keyword banners should closely monitor internet users’ use of search keywords that are relevant for their business by using data provided by KeywordCity.com or WordSpot.com, for example. Once commonly used keywords are identified, the company could then classify them into different sets based on hyponym relationships. For instance, a travel website may discover that “world travel” and “budget travel” are commonly used hyponyms of “travel.” Then it can develop two specific banners that correspond to these two keywords. Suppose that the advertiser further uncovers more specific keywords that are hyponyms of those two words. Then it can develop even more specific banners accordingly. This process and sample banners are illustrated in Table 3. By developing multi-version banners for the search engine, the advertiser can stand a better chance in getting accurate banners that can catch consumer attention and motivate visits to its website.

FURTHER STUDY

The present study is based on a sample of keywords composed from major search directories. Future studies can extend our research in several directions. For example, a more comprehensive study may rely on common keywords provided by a probabilistic sample of internet users. However, in that scenario the respondents must be able to generate sets of keywords for each category that only differ in the narrowness of range. Given that different consumers may have different understandings of secondary or tertiary levels of keywords for a broad category (e.g., travel), there may be inconsistencies among users in deciding how much “distance” there is between a “broad” and “moderate” keyword and between a “moderate” and “narrow” keyword. With such inconsistencies, the classification of keywords into three different ranges may not be meaningful and comparable across different respondents.

Second, this study used the popularity of search engines as a proxy for the amount of banner ad inventories. Based on this assumption, we studied popular search engines, supposedly with bigger ad inventories, and their technical capability to produce more precise banner matches. However, it would be worthwhile to get access to the number and types of banner inventories available for keywords in top search engines to confirm this assumption. With such information, future research can also test whether banner ad copy inventories are positively related to the number of exact banner matches.

Given the growing popularity of keyword-activated banners, online media companies must identify factors that may affect the performance of smart banners. Future research is certainly needed to further our understanding of this phenomenon and help online advertisers make more informed decisions.

TABLE 3

A Sample Multiversion Keyword Banner Strategy for a Travel Website

<table>
<thead>
<tr>
<th>Search Keywords and Types</th>
<th>Corresponding Banners</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>World Travel</strong></td>
<td></td>
</tr>
<tr>
<td>(“Moderate”)</td>
<td>Travel to Europe (“Narrow”)</td>
</tr>
<tr>
<td></td>
<td>Fall Savings on Travel to Europe!</td>
</tr>
<tr>
<td></td>
<td>Experience the Best of Asia!</td>
</tr>
<tr>
<td><strong>Travel</strong></td>
<td></td>
</tr>
<tr>
<td>(“broad”)</td>
<td>Lowest Fares (“Narrow”)</td>
</tr>
<tr>
<td></td>
<td>Lowest Fares Guaranteed!</td>
</tr>
<tr>
<td><strong>Budget Travel</strong></td>
<td></td>
</tr>
<tr>
<td>(“Moderate”)</td>
<td>Budget hotels (“Narrow”)</td>
</tr>
<tr>
<td></td>
<td>Heavy Discounts on Hotel Stays!</td>
</tr>
</tbody>
</table>

**WEB YU DOU** is assistant professor of marketing at St. Cloud State University. His previous works have been published in the *Journal of Advertising Research*, *Industrial Marketing Management*, and the *Quarterly Journal of Electronic Commerce*. His main research interests include marketing communications in the cyber world and global electronic commerce.

42 JOURNAL OF ADVERTISING RESEARCH July • August 2001
Randy Lim is manager at Wells Fargo Bank in Sauk Rapids, Minnesota. He received his MBA degree from St. Cloud State University. His research interest lies primarily in financial applications in e-commerce.

Susan Yang is associate director of institutional research at St. Cloud State University. She obtained her MBA degree from Marquette University. Her research interests include management information systems for higher institutions and web-based programming.

REFERENCES


